Analysis of CMB polarization on an incomplete sky

Antony Lewis,[1,*] Anthony Challinor,[2,†] and Neil Turok[1,‡]

[1]*DAMTP, CMS, Wilberforce Road, Cambridge CB3 0WA, UK.*

[2]*Astrophysics Group, Cavendish Laboratory, Madingley Road, Cambridge CB3 OHE, UK.*

The full sky cosmic microwave background polarization field can be decomposed into 'electric' and 'magnetic' components. Working in harmonic space we construct window functions that allow clean separation of the electric and magnetic modes from observations over only a portion of the sky. We explicitly demonstrate the method for azimuthally symmetric patches, but also present it in a form in principle applicable to arbitrarily-shaped patches. From the window functions we obtain variables that allow for robust estimation of the magnetic component without risk of contamination from the probably much larger electric signal. The variables have a very simple noise properties, and further analysis using them should be no harder than analysing the temperature field. For an azimuthally-symmetric patch, such as that obtained from survey missions when the galactic region is removed, the exactly-separated variables are fast to compute. We estimate the magnetic signal that could be detected by the Planck satellite in the absence of extra-galactic foregrounds. We also discuss the sensitivity of future experiments to tensor modes in the presence of a magnetic signal generated by weak lensing, and give lossless methods for analysing the electric polarization field in the case that the magnetic component is negligible. A series of appendices review the spin weight formalism and give recursion relations for fast computation of the spin-weighted spherical harmonics and their inner products over azimuthally-symmetric patches of the sphere. A further appendix discusses the statistics of weak signal detection.

I. INTRODUCTION

Observations of fluctuations in the temperature of the cosmic microwave background (CMB) are now providing us with a direct view of the primordial inhomogeneities in the universe. The power spectrum of temperature fluctuations yields a wealth of information on the nature of the primordial perturbations, and the values of the cosmological parameters. Mapping the polarization of the cosmic microwave sky is an important next step, offering a great deal of complementary information, especially regarding the character of the primordial inhomogeneities [1]. One of the most interesting questions to resolve is whether the primordial perturbations possessed a tensor (gravitational wave) component, as predicted by simple inflationary models. Here, polarization measurements offer a unique probe [2, 3, 4].

Polarization of the cosmic microwave sky is produced by electron scattering, as photons decouple from the primordial plasma. Linear polarization is produced when there is a quadrupole component to the flux of photons incident on a scattering electron. Scalar (density) perturbations generate an 'electric' (gradient) polarization pattern on the sky due to gradients in the velocity field on the surface of last scattering. For scalar perturbations the velocity field is curl-free, and this leads directly to the production of an entirely 'electric' pattern of linear polarization. In contrast, tensor perturbations (gravitational waves) produce polarization by anisotropic redshifting of the energy of photons through decoupling. In this case the polarization has 'magnetic' (i.e. curl) and 'electric' (i.e. gradient) components at a comparable level. A magnetic signal can also be produced by weak lensing of the electric polarization generated by scalar modes. Detection and analysis of the lensing signal would be interesting in itself, but a detection of an additional tensor component would provide strong evidence for the presence of primordial gravitational waves, a generic signature of simple inflationary models.

Detecting or excluding a magnetic component is clearly of fundamental significance in cosmology. But there is a significant obstacle to be faced. The problem is that for the foreseeable future, the primordial sky polarization will only be observable over the region of the sky which is not contaminated by emission from our galaxy and other foreground sources of polarization. Thus we shall only be able to measure the polarization over a fraction of the sky. But the electric/magnetic decomposition is inherently *non-local*, and *non-unique* in the presence of boundaries.

*Electronic address: Antony@AntonyLewis.com

†Electronic address: A.D.Challinor@mrao.cam.ac.uk

‡Electronic address: N.G.Turok@damtp.cam.ac.uk

To understand this, consider the analogous problem of representing a vector field V_i (in two dimensions) as a gradient plus a curl:

$$V_i = \nabla_i \Phi + \epsilon_i{}^j \nabla_j \chi, \tag{1}$$

the electric and magnetic components respectively. From this equation, one has $\nabla^2\Phi = \nabla^i V_i$, and $\nabla^2\chi = -\nabla^i \epsilon_i{}^j V_j$. For a manifold without a boundary, like the full sky, the Laplacian may be inverted up to a constant zero mode, and the two contributions to V_i are uniquely determined. But for a finite patch, one can always think of adding charged sources for the potentials Φ and χ outside of the patch on which V_i is measured, which alter Φ and χ without changing V_i. For example one can add to Φ and χ pieces with equal but perpendicular gradients so there is no net contribution to V_i.

Since full sky observations are unrealistic, so is the hope of a unique decomposition of the sky polarization into electric and magnetic components. However, this does not at all mean that the hypothesis of a magnetic signal cannot be tested. One possibility is to construct a local measure of the magnetic signal by differentiating the measured polarization (i.e. $\nabla^i \epsilon_i{}^j V_j$ vanishes if V_i is pure electric in the analogue example above), but this is problematic for noisy, sampled data. A more promising alternative, which avoids differentiating the data, is to construct line integrals of the polarization [5, 6]. For example, in the vector analogy above, any line integral $\oint \mathrm{d}x^i V_i$ is guaranteed to vanish if V_i is purely electric. However, the problem with these line integrals is that there are an infinite number of them, and they are not statistically independent. One would therefore prefer a set of 'magnetic' variables to which the 'electric' component does not contribute, but which are finite in number and statistically independent, for a rotationally symmetric statistical ensemble. Since polarization from a primordial scale invariant spectrum of gravitational waves is predominantly generated on scales of a degree or so (the angle subtended by the horizon at last scattering), we would expect to be able to characterize the cosmic magnetic signal by a set of statistically independent variables roughly equal in number to the size of the patch in square degrees. However the signal within a degree or so of the boundary cannot be unambiguously characterized as magnetic, and hence one loses a number of variables proportional to the length of the boundary. The amount of information about the magnetic signal therefore increases as the patch area minus the area of this boundary layer. In this paper we shall find the set of observable 'magnetic' variables explicitly for circular sky patches: the method may be generalized to non-circular patches if necessary.

As mentioned above, the electric component of the polarization (due primarily to scalar perturbations) is expected to be much larger than the magnetic signal. Therefore to detect the latter it may be useful to construct observables which suffer no electric contamination. We show how to construct such variables, and use them to estimate what magnitude of magnetic signal the planned Planck satellite[1] might be able to detect. We also discuss the optimal survey size for future experiments aimed at detecting tensor modes via magnetic polarization, including the effects of 'magnetic noise' due to weak lensing of the dominant electric polarization [7, 8]. Even for observations that do not expect to detect the magnetic signal the magnetic-only observables are likely to be very useful in checking consistency of any residual polarization with noise or indeed in identifying foreground contamination. They may also be useful for studying the small scale weak lensing signal [9].

To construct variables that depend only on the electric or magnetic polarization we integrate the polarization field over the observed patch with carefully chosen spin-weight 2 window functions. We present a harmonic-based approach for constructing these window functions which is exact in the limit of azimuthally-symmetric patches. The method is expected still to perform well for arbitrary shaped patches of the sky, but the separation will no longer be exact in that case. Constructing the window functions with our harmonic method automatically removes redundancy due to the finite size of the patch, keeps the information loss small (except for very small patches), and ensures that for idealized noise in the polarization map (isotropic and uncorrelated), the noise on the electric and magnetic variables preserves these properties. In this respect the construction is analogous to the orthogonalized harmonics approach used in the analysis of temperature anisotropies [10, 11]. However in the polarized case there is no simple interpretation in terms of a set of orthogonalized harmonics.

In Ref. [12] it was shown how the lossless quadratic estimator technique can be applied to polarization. There, no attempt was made to separate the electric and magnetic contribution to the estimators, so the resulting window functions for the power displayed considerable leakage between the electric and magnetic modes. The authors of Ref. [12] showed how the leakage could be reduced, but it is arguably still too large to allow robust estimation of the magnetic signal in the presence of an electric signal that is orders of magnitude larger. We are able to perform a much cleaner separation at the level of the harmonic components in the map, and as we shall see the information loss in our approach is quite small for full sky surveys with a galactic cut.

---

[1] http://astro.estec.esa.nl/Planck

The electric-magnetic decomposition of the polarization field is exactly analogous to the corresponding decomposition of projected galaxy ellipticities induced by weak lensing [13, 14]. Ref. [15] shows how to construct local real-space correlation functions for measuring the magnetic component. These are useful for distinguishing the purely electric signal due to gravitational lensing from intrinsic correlations in galaxy alignments, and the method has the advantage of working for arbitrarily shaped regions of sky. However the method assumed a flat sky approximation, and includes only the two-point information. For polarization observations the sky curvature will be important and we aim to extract a set of statistically independent observables that contain as much of the magnetic information as possible. This may also prove useful for weak lensing studies.

The paper is arranged as follows. In Sec. II we present the spin-weight 2 window technique for separating electric and magnetic polarization on the sphere, generalizing results in Refs. [5, 6]. Section III describes our harmonic-based technique for constructing window functions with the properties required to ensure separation of the electric and magnetic modes while keeping information loss small. Classical techniques for testing the hypothesis that there is no magnetic signal are discussed in Sec. IV, and estimates of the detection limits with the Planck satellite and future experiments are also given. Lossless methods for estimation of the polarization power spectra are contrasted with methods using the separated variables in Sec. V. In a series of appendices we outline our conventions for spin weight functions and their spherical harmonics. In addition we present a number of the standard integral theorems on 2-dimensional manifolds in convenient spin weight form, and present recursive methods for the fast computation of the spin weight spherical harmonics and their inner products over azimuthally symmetric patches of the sphere. A further appendix discusses the statistics of detecting weak signals from tensor modes.

II. ELECTRIC AND MAGNETIC POLARIZATION

The observable polarization field is described in terms of the two Stokes' parameters Q and U with respect to a particular choice of axes about each direction on the sky. In this paper we take these axes to form a right-handed set with the incoming radiation direction (following Ref. [3]). The real Stokes' parameters are conveniently combined into a single complex field that represents the observed polarization

$$P = Q + iU. \tag{2}$$

The values of Stokes' parameters depend on the choice of axes; since Q is the difference of the intensity in two orthogonal directions it changes sign under a rotation of 90°. The Q field is related to the U field by a rotation of 45°. More generally under a right-handed rotation of the axes by an angle α about the incoming direction the complex polarization transforms as $P \to e^{-2i\alpha}P$ and is therefore described as having spin minus two (see Appendix A for our conventions for spin weight functions). The analysis of polarized data is therefore rather more complicated than for the temperature which does not depend on making a choice of axes in each direction on the sky.

As described in Appendix A, one can define spin raising and lowering operators that can be used to relate quantities of different spin [16, 17]. The spin raising operator is denoted $\eth$ and the lowering operator $\overline{\eth}$. Since the polarization has spin-weight -2 it can be written as the action of two spin lowering operators on a spin zero complex number

$$P = \overline{\eth}\,\overline{\eth}\,(P_E + iP_B). \tag{3}$$

The underlying real scalar (spin-zero) fields P_E and P_B describe electric and magnetic polarization respectively [17]. They are clearly non-local functions of the Stokes' parameters. One can define a spin zero quantity which is local in the polarization by acting with two spin raising operators. Using some results from Appendix A one obtains

$$\eth\,\eth\,P = (\nabla^2 + 2)\nabla^2(P_E + iP_B) \tag{4}$$

where ∇ is the covariant derivative on the sphere. The real and imaginary parts of this equation can therefore be used to determine the electric and magnetic parts of the polarization. Performing a surface integral we define

$$I'_{-2,W} \equiv \int_S \mathrm{d}S\, W^*\,\eth\,\eth\,P, \qquad I'_{2,W} \equiv \int_S \mathrm{d}S\, W^*\,\overline{\eth}\,\overline{\eth}\,P^*, \tag{5}$$

where W is a complex window function defined over some patch S of the observed portion of the sky. It follows that

$$E'_W \equiv \tfrac{1}{2}(I'_{2,W} + I'_{-2,W}), \qquad B'_W \equiv -i\tfrac{1}{2}(I'_{2,W} - I'_{-2,W}) \tag{6}$$

provide a measure of the electric and magnetic signals. Note that $E'_{W^*} = E'^*_W$ with an equivalent result for B'_W. Using the integral theorem (A23) in Appendix A we can write

$$I'_{2,W} = \int_S \mathrm{d}S\, P^* {}_2W^* + \oint_{\partial S} {}_1\overline{\mathrm{d}l}\, \left(W^* \overline{\eth} P^* - P^* \overline{\eth} W^*\right), \tag{7}$$

$$I'_{-2,W} = \int_S \mathrm{d}S\, P {}_{-2}W^* + \oint_{\partial S} {}_1\mathrm{d}l\, \left(W^* \eth P - P\, \eth W^*\right), \tag{8}$$

where ${}_2W \equiv \eth\, \eth W$ is now a spin 2 window function, ${}_{-2}W \equiv \overline{\eth}\, \overline{\eth} W$ is a spin -2 window function, and ${}_1\mathrm{d}l = {}_1\overline{\mathrm{d}l}^*$ is the spin 1 element of length around the boundary ∂S of S. Clearly we do not wish to take derivatives of noisy observed data and hence it is usually useful to choose the window function to eliminate the derivative terms on the boundary.

For CMB polarimetry we are interested in the polarization defined on the spherical sky. The surface integrals vanish if we choose W such that $\eth\, \eth W = \overline{\eth}\, \overline{\eth} W = 0$, which will be true if W is a linear combination of the spherical harmonics with $l = 0$ or 1, since these possess no spin 2 component. If we then set $W = 0$ on the boundary, so as to eliminate the derivatives of the polarization, we are forced to consider circular patches S, in which case a combination of the two $m = 0$ harmonics works. This implies that the electric and magnetic signals can be probed by performing line integrals around circles, as emphasized in Refs. [5, 6]. These line integrals can be performed around any circle that is enclosed in the observed region of the sky, and it is unclear how to obtain a complete set of statistically independent observables in order to extract all of the available information. Also for current experiments, performing one-dimensional line integrals on pixelized maps is unlikely to be a good way to extract information robustly.

In this paper, we suggest choosing the window functions so that the line integrals around ∂S that appear in the construction of E'_W and B'_W contain no contribution from the magnetic and electric polarization respectively. In the absence of special symmetries (see below for exceptions that arise in the case of circular patches) this requires that W, $\eth W$, and $\overline{\eth} W$ all vanish on the boundary. These conditions are equivalent to demanding that the window function and its normal derivative vanish on ∂S. With such a choice of window we can measure the electric and magnetic signals using only the surface integrals

$$I_{\pm 2,W} \equiv E_W \pm iB_W \equiv \int_S \mathrm{d}S\ {}_{\pm 2}W^*(Q \mp iU). \tag{9}$$

Since the window functions are scalar functions on the sphere we can expand them in spherical harmonics,

$$W = \sum_{l\geq 2} \sum_{|m|\leq l} \sqrt{\frac{(l-2)!}{(l+2)!}} W_{lm} Y_{lm}. \tag{10}$$

(The square root factor is included for later convenience.) We need not include $l = 0$ and 1 spherical harmonics since they do not contribute to the spin-weight ± 2 window functions, and the boundary integral terms automatically separate for these multipoles. In practice, we are only interested in probing scales to some particular $l_{\max}$ (e.g. the magnetic signal from tensor modes has maximal power for $l \approx 100$ and decreases rapidly with l), so the sum in Eq. (10) can be truncated at some finite $l_{\max}$.

We shall focus on the case where the observed sky patch is azimuthally symmetric in which case the construction of exact window functions becomes particularly simple. The harmonic-based method we describe in Sec. III provides a practical solution to constructing a non-redundant set of window functions that separate the electric and magnetic modes exactly. In addition, for the special case of isotropic, uncorrelated noise on the observed polarization, these simple properties are preserved in the variables E_W and B_W. For observations over non-azimuthally symmetric patches our method can, of course, be used over the largest inscribed circular patch, but in this case there is inevitable information loss since we use only a subset of the observed data. However, we expect that the method presented in Sec. III could also be applied directly to the full observed region to construct window functions that achieve approximate separation of electric and magnetic polarization.

Consider the case of an azimuthally-symmetric patch so the boundary ∂S consists of one or two small circles. For each azimuthal dependence on m we can construct combinations

$$W_m = \sum_{l=\max(2,|m|)}^{l_{\max}} \sqrt{\frac{(l-2)!}{(l+2)!}} W_{lm} Y_{lm} \tag{11}$$

that satisfy the necessary boundary conditions. For $m = 0$ it is easy to see that E_W and B_W contain no contribution from P_B and P_E respectively for any choice of the W_{l0} [i.e. the boundary integrals that distinguish E_W (B_W) from

E'_W (B'_W) vanish if the polarization is pure magnetic (electric)]. It follows that for $m = 0$ there are $l_{\rm max} - 1$ linearly independent window functions that satisfy the boundary conditions. For $|m| = 1$ it will be shown in the next section that there is only one independent linear constraint per boundary circle, so there are $l_{\rm max}-2$ possible window functions ($l_{\rm max} - 3$ for a boundary composed of two circles). For $|m| \geq 2$ there are two linear constraints per boundary circle which can be taken to be the vanishing of W_m and its normal derivative. In this case there are $l_{\rm max} - |m| - 1$ ($l_{\rm max} - |m| - 3$) window functions for boundaries consisting of one (two) small circles.

Since we are only considering a fraction of the sky not all of the window functions counted above may return observables E_W and B_W containing independent information. This arises because for large $l_{\rm max}$, or small patches, there will generally arise non-zero window coefficients W_{lm} that produce spin 2 window functions that are poorly supported over the patch. (See e.g. Ref. [11] for a discussion of the equivalent problem in the case of scalar functions.) The redundancy in the set of acceptable window functions can be removed by expanding the spin 2 window functions in a smaller set of functions which are (almost) complete for band-limited signals over the region S. The construction of such a set by singular value methods (e.g. Refs. [11, 18]) forms the starting point of the method we present in the Sec. III.

A. Harmonic expansion

We construct window functions in harmonic space, so as a useful preliminary we consider the harmonic expansion of spin-weight 2 fields over the full sphere [3, 4]. The polarization $P \equiv Q + iU$ is spin -2 and can be expanded over the whole sky in terms of the spin two harmonics (see Appendix B for our conventions and some useful results)

$$Q \pm iU = \sum_{lm} a_{\mp 2,lm}\, {}_{\mp 2}Y_{lm} = \sum_{lm} (E_{lm} \mp iB_{lm})_{\mp 2}Y_{lm}. \tag{12}$$

Reality of Q and U requires $a^*_{-2,lm} = (-1)^m a_{2,l(-m)}$, so that $E^*_{lm} = (-1)^m E_{l(-m)}$ with an equivalent result for B_{lm}. Under parity transformations $E_{lm} \to (-1)^l E_{lm}$ but $B_{lm} \to (-1)^{l+1} B_{lm}$, since ${}_sY_{lm}(\pi - \theta, \phi + \pi) = (-1)^l {}_{-s}Y_{lm}(\theta, \phi)$. From the orthogonality of the spherical harmonics over the full sphere it follows that

$$E_{lm} = \tfrac{1}{2}(a_{2,lm} + a_{-2,lm}) = \tfrac{1}{2}\int_{4\pi} {\rm d}S\, P\, {}_{-2}Y^*_{lm} + \tfrac{1}{2}\int_{4\pi} {\rm d}S\, P^*\, {}_2Y^*_{lm}, \tag{13}$$

$$iB_{lm} = \tfrac{1}{2}(a_{2,lm} - a_{-2,lm}) = -\tfrac{1}{2}\int_{4\pi} {\rm d}S\, P\, {}_{-2}Y^*_{lm} + \tfrac{1}{2}\int_{4\pi} {\rm d}S\, P^*\, {}_2Y^*_{lm}. \tag{14}$$

In a rotationally-invariant ensemble, the expectation values of the harmonic coefficients define the electric and magnetic polarization power spectra:

$$\langle E^*_{l'm'} E_{lm}\rangle = \delta_{l'l}\delta_{m'm} C_l^{EE}, \qquad \langle B^*_{l'm'} B_{lm}\rangle = \delta_{l'l}\delta_{m'm} C_l^{BB}. \tag{15}$$

If the ensemble is parity-symmetric the cross term is zero, $\langle E^*_{l'm'} B_{lm}\rangle = 0$.

The form of the harmonic expansion (10) of the window function ensures that the spin-weight ± 2 windows are

$${}_{\pm 2}W = \sum_{lm} W_{lm}\, {}_{\pm 2}Y_{lm}, \tag{16}$$

where the sum is over $l \geq 2$ and $|m| \leq l$. Evaluating the surface integrals in Eq. (9) we find

$$E_W = \sum_{lm} W^*_{lm} \tilde{E}_{lm}, \qquad B_W = \sum_{lm} W^*_{lm} \tilde{B}_{lm}, \tag{17}$$

where the pseudo-harmonics are obtained by restricting the integrals in Eqs. (13) and (14) to the region S:

$$\tilde{E}_{lm} = \frac{1}{2}\sum_{l'm'} \int_S {\rm d}S\, \left[(E_{l'm'} - iB_{l'm'})\, {}_{-2}Y_{l'm'}\, {}_{-2}Y^*_{lm} + (E_{l'm'} + iB_{l'm'})\, {}_2Y_{l'm'}\, {}_2Y^*_{lm}\right], \tag{18}$$

$$\tilde{B}_{lm} = \frac{i}{2}\sum_{l'm'} \int_S {\rm d}S\, \left[(E_{l'm'} - iB_{l'm'})\, {}_{-2}Y_{l'm'}\, {}_{-2}Y^*_{lm} - (E_{l'm'} + iB_{l'm'})\, {}_2Y_{l'm'}\, {}_2Y^*_{lm}\right]. \tag{19}$$

Defining Hermitian coupling matrices

$$W_{\pm(lm)(lm)'} \equiv \tfrac{1}{2}({}_2W_{(lm)(lm)'} \pm {}_{-2}W_{(lm)(lm)'}), \tag{20}$$

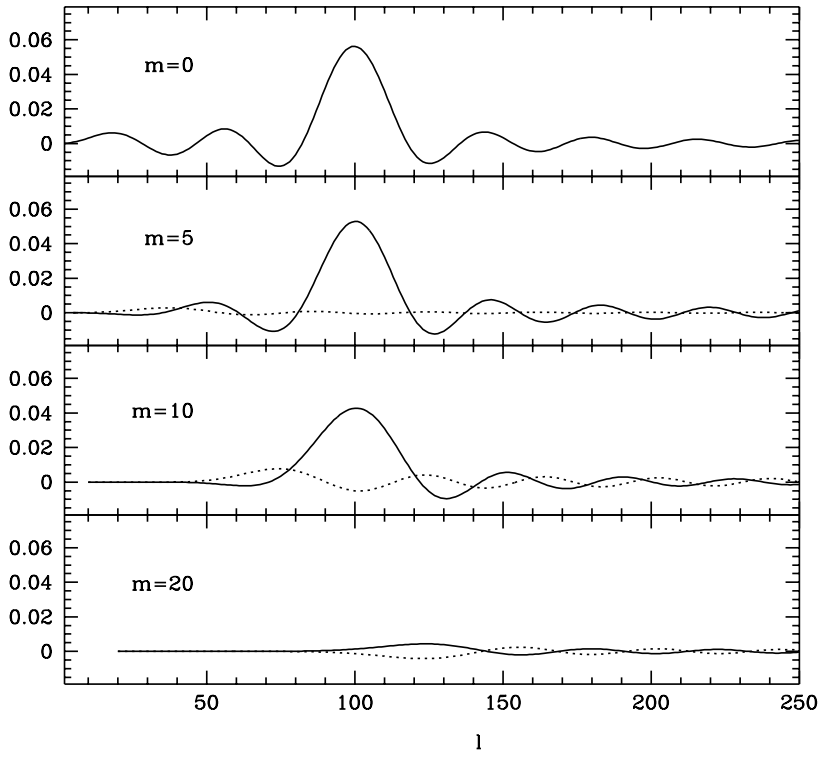

FIG. 1: The window functions $W_{+(l'm)(lm)}$ (solid lines) and $W_{-(l'm)(lm)}$ (dashed lines) for $l' = 100$ and various m for an azimuthally symmetric patch with $\theta < 10^\circ$. The dashed lines show the E_{lm} contamination of $\tilde{B}_{l'm}$ as a function of l. For $m = 0$ there is no contamination, and as m increases the functions decrease in amplitude as the corresponding harmonics become more localized outside of the patch.

where

$$ {}_sW_{(lm)(lm)'} \equiv \int_S \mathrm{d}S\, {}_sY^*_{lm}\, {}_sY_{l'm'}, \tag{21} $$

we can write

$$ \tilde{E}_{lm} = \sum_{l'm'} \left(W_{+(lm)(lm)'} E_{l'm'} + i W_{-(lm)(lm)'} B_{l'm'} \right), \tag{22} $$

$$ \tilde{B}_{lm} = \sum_{l'm'} \left(W_{+(lm)(lm)'} B_{l'm'} - i W_{-(lm)(lm)'} E_{l'm'} \right). \tag{23} $$

In the limit $l_{\max} \to \infty$, ${}_{\pm2}W_{(lm)(lm)'}$ become projection operators as a consequence of the completeness of the spin-weight harmonics. The matrix $W_{-(lm)(lm)'}$ controls the contamination of E_W and B_W with magnetic and electric polarization respectively. Our aim is to construct window functions W_{lm} that remove this contamination for all E_{lm} and B_{lm}. Some elements of the matrices $W_{\pm(lm)(lm)'}$ are shown in Fig. 1.

For azimuthally-symmetric patches the coupling matrices are block diagonal ($W_{\pm(lm)(lm)'} \propto \delta_{mm'}$), and so window functions can be constructed for each m separately [see Eq. (11)]. For $m = 0$ we have

$$ \int_S \mathrm{d}S_2 Y_{l'0}\, {}_2Y^*_{l0} = \int_S \mathrm{d}S\, {}_{-2}Y_{l'0}\, {}_{-2}Y^*_{l0}, \tag{24} $$

so $W_{-(l0)(l'0)} = 0$ and we have clean separation for any azimuthally-symmetric window function. The set of azimuthally symmetric window functions gives $l_{\rm max}-1$ separated variables that contain the same information as would be obtained by computing line integrals around all those circles concentric with the boundary of the azimuthal patch. For general m there is leakage of E_{lm} into $\tilde{B}_{lm}$; for parity-symmetric cuts there is only leakage between modes with different parity (i.e. for even l the pseudo-harmonics $\tilde{B}_{lm}$ depend on $E_{l'm}$ only for odd l').

We showed in the previous section that, for a general window function, the contamination of e.g. E_W by the magnetic polarization is due entirely to boundary terms. This implies that $W_{-(lm)(lm)'}$ can always be written as a line integral around the boundary of S. (We show in Appendix C that the matrices ${}_sY_{(lm)(lm)'}$ can be transformed into line integrals for $l \neq l'$. However $W_{-(lm)(lm)'}$ can be written as a line integral for all l and l'.) Making use of Eq. (C9), it is straightforward to show that

$$\begin{aligned} W_{-(lm)(lm)'} &= \frac{1}{2}\sqrt{\frac{(l-2)!}{(l+2)!}}\Big(\oint_{\partial S} {}_1\overline{\mathrm{d}l}\ \left[\sqrt{l(l+1)}{}_1Y^*_{lm}{}_2Y_{l'm'} + \sqrt{(l'-1)(l'+2)}Y^*_{lm}{}_1Y_{l'm'}\right] \\ &\quad + \oint_{\partial S} {}_1\mathrm{d}l\ \left[\sqrt{l(l+1)}{}_{-1}Y^*_{lm}{}_{-2}Y_{l'm'} + \sqrt{(l'-1)(l'+2)}Y^*_{lm}{}_{-1}Y_{l'm'}\right]\Big). \end{aligned} \tag{25}$$

This can be put in manifestly Hermitian form using the recursion relation derived from the action of $(\eth - \overline{\eth})$ on ${}_sY_{lm}$

$$\frac{m + s\cos\theta}{\sin\theta}{}_sY_{lm} = \tfrac{1}{2}\sqrt{(l-s)(l+s+1)}{}_{s+1}Y_{lm} + \tfrac{1}{2}\sqrt{(l+s)(l-s+1)}{}_{s-1}Y_{lm}. \tag{26}$$

For a circular boundary at constant latitude $\theta = \Theta$ (i.e. the boundary of an azimuthal patch), we find

$$W_{-(lm)(l'm)} = -4m\pi[u_l(m)u^*_{l'}(m) + v_l(m)v^*_{l'}(m)], \tag{27}$$

where the vectors

$$u_l(m) = \sqrt{\frac{(l-2)!}{(l+2)!}}\left[\partial_\theta Y_{lm}(\Theta,\Phi) - \cot\Theta Y_{lm}(\Theta,\Phi)\right], \tag{28}$$

$$v_l(m) = \sqrt{\frac{(l-2)!}{(l+2)!}}\frac{\sqrt{(m^2-1)}}{\sin\Theta}Y_{lm}(\Theta,\Phi) \tag{29}$$

for $l \geq 2$ and some arbitrary Φ. [Note that $u_l(m)$ and $v_l(m)$ will not generally be orthogonal so Eq. (27) is not the spectral decomposition of $W_{-(lm)(l'm)}$.] Any window W_{lm} whose inner products with $u_l(m)$ and $v_l(m)$ both vanish, i.e.

$$\sum_l W^*_{lm}u_l(m) = \sum_l W^*_{lm}v_l(m) = 0, \tag{30}$$

will achieve clean separation of electric and magnetic polarization. For $|m| > 1$ such window functions and their normal derivative necessarily vanish on the boundary. As noted earlier, for $|m| = 1$ there is actually only one constraint to be satisfied which now follows from the fact that $v_l(\pm 1) = 0$.

III. ELECTRIC/MAGNETIC SEPARATION ON THE CUT SKY

In this section we give a practical method for constructing a non-redundant set of window functions $\{W_I\}$ where I labels the particular window, that achieve exact separation for azimuthal patches. The corresponding (cleanly separated) electric and magnetic observables will be denoted E_{W_I} and B_{W_I}. We will make use of a notation where vectors are denoted by bold Roman font, e.g. $\mathbf{B}_W$ has components B_{W_I}, and $\mathbf{B}$ has components B_{lm}, and matrices are denoted by bold italic font, e.g. $\boldsymbol{W}_\pm$ have components $W_{\pm(lm)(lm)'}$. We present the method in a form that is applicable (though no longer exact) to arbitrary shaped regions S; for azimuthal patches the method is exact. For the azimuthal case all matrices are block diagonal and the window functions can be constructed for each m separately.

In matrix form, Eq. (17) is

$$\mathbf{E}_W = \boldsymbol{W}^*\tilde{\mathbf{E}}, \qquad \mathbf{B}_W = \boldsymbol{W}^*\tilde{\mathbf{B}}, \tag{31}$$

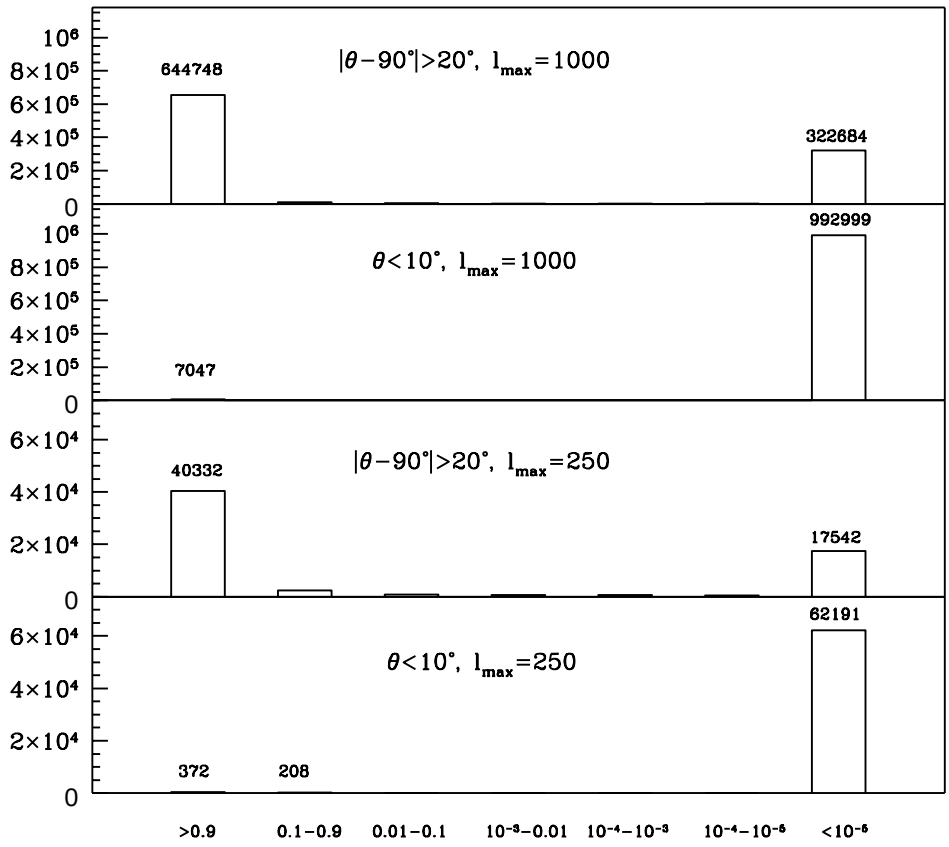

FIG. 2: The distribution of eigenvalues of $\boldsymbol{W}_+$ for two azimuthally-symmetric sky cuts with $l_{\rm max} = \{250, 1000\}$. The distribution is approximately bimodal, and the fraction of the eigenvalues corresponding to well-determined modes (eigenvalue significantly non-zero) is given by the fraction of the sky area in the patch in the limit $l_{\rm max} \to \infty$.

where $\boldsymbol{W} = W_{I(lm)}$ is the matrix whose Ith row contains the harmonic coefficients of the Ith window function, and recall

$$\tilde{\mathbf{E}} = \boldsymbol{W}_+\mathbf{E} + i\boldsymbol{W}_-\mathbf{B}, \qquad \tilde{\mathbf{B}} = \boldsymbol{W}_+\mathbf{B} - i\boldsymbol{W}_-\mathbf{E}. \tag{32}$$

For an azimuthally-symmetric sky patch the block-diagonal matrices ${}_{\pm2}\boldsymbol{W}$ (with components ${}_{\pm2}W_{(lm)(l'm)}\delta_{mm'}$) from which $\boldsymbol{W}_\pm$ are constructed can be computed very quickly using the recursion relations given in Appendix C. Alternatively, $\boldsymbol{W}_-$ can be computed directly from Eq. (27). In the limit of full sky coverage $\boldsymbol{W}_- \to 0$ and $\boldsymbol{W}_+ \to \boldsymbol{I}$. We know that for $|m| \geq 2$, the range of mth submatrix of $\boldsymbol{W}_-$ is two-dimensional [spanned by $u_l(m)$ and $v_l(m)$], so that all but two of the eigenvalue of the submatrix are zero. Equivalently, all but two linear combinations of the $\tilde{B}_{lm}$ are independent of $E_{l'm}$. The $|m| = 1$ submatrices of $\boldsymbol{W}_-$ have only one non-zero eigenvalue; the associated eigenvectors are $u_l(\pm1)$. The $m = 0$ submatrix is identically zero. The essence of our method for constructing the window functions is to choose $\boldsymbol{W}$ to project out of the range of $\boldsymbol{W}_-$.

We first diagonalize $\boldsymbol{W}_+ = \boldsymbol{U}_+\boldsymbol{D}_+\boldsymbol{U}_+^\dagger$ by performing a singular value decomposition [18]. Here, $\boldsymbol{D}_+$ is a positive (semi-)definite diagonal matrix whose elements are the eigenvalues of $\boldsymbol{W}_+$. The columns of the unitary matrix $\boldsymbol{U}_+$ are the normalized eigenvectors of $\boldsymbol{W}_+$. The singular value decomposition allows us to identify the linear combinations $\boldsymbol{U}_+^\dagger\mathbf{B}$ that are poorly determined by $\tilde{\mathbf{B}}$ — those corresponding to the small diagonal elements of $\boldsymbol{D}_+$. The eigenvectors

with very small eigenvalues correspond to polarization patterns that essentially have no support inside the observed patch of the sky, and would lead to a set of redundant window functions if not removed from the analysis. The distribution of eigenvalues of $\boldsymbol{W}_+$ is approximately bimodal as illustrated in Fig. 2, and the exact definition of 'small' is not critical when considering the range of the matrix. This bimodality arises because ${}_{\pm 2}\boldsymbol{W}$ are approximately projection operators for large $l_{\rm max}$, and the fact that the range of $\boldsymbol{W}_-$ is a rather small subspace. To remove redundant degrees of freedom from the spin 2 window functions, we define an operator $\tilde{\boldsymbol{U}}_+$ which projects onto the eigenvectors of $\boldsymbol{W}_+$ whose eigenvalues are close to one. This amounts to removing the appropriate columns of $\boldsymbol{U}_+$. Since $\boldsymbol{U}_+$ is orthogonal, $\tilde{\boldsymbol{U}}_+$ is column orthogonal and hence $\tilde{\boldsymbol{U}}_+^\dagger \tilde{\boldsymbol{U}}_+ = \boldsymbol{I}$ (but $\tilde{\boldsymbol{U}}_+ \tilde{\boldsymbol{U}}_+^\dagger \neq \boldsymbol{I}$). The matrix $\tilde{\boldsymbol{D}}_+$ is the corresponding smaller square diagonal matrix, and we have

$$\tilde{\boldsymbol{U}}_+^\dagger \tilde{\mathbf{B}} \approx \tilde{\boldsymbol{D}}_+ \tilde{\boldsymbol{U}}_+^\dagger \mathbf{B} - i \tilde{\boldsymbol{U}}_+^\dagger \boldsymbol{W}_- \mathbf{E}. \tag{33}$$

We now multiply by $\tilde{\boldsymbol{D}}_+^{-1/2}$, defined by $[\tilde{\boldsymbol{D}}_+^{-1/2}]_{ij} \equiv \delta_{ij} [\tilde{\boldsymbol{D}}_+]_{ii}^{-1/2}$, to give

$$\tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger \tilde{\mathbf{B}} \approx \tilde{\boldsymbol{D}}_+^{1/2} \tilde{\boldsymbol{U}}_+^\dagger \mathbf{B} - i \tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger \boldsymbol{W}_- \mathbf{E}. \tag{34}$$

Including the factor $\tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger$ in the window functions $\boldsymbol{W}$ is equivalent to constructing the spin 2 window functions from a reduced basis that is orthonormal and (almost) complete over the region S.

The remaining step is to project out the unwanted boundary terms that contaminate $\tilde{\mathbf{B}}$ with $\mathbf{E}$. For the case of an azimuthal patch, the ranges of the submatrices of $\boldsymbol{W}_-$ are all two dimensional (or lower). It follows that the same is true of $\tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger \boldsymbol{W}_-$, so we can remove the unwanted boundary term by ensuring that $\boldsymbol{W}$ also includes a factor that projects out of the range of $\tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger \boldsymbol{W}_-$. In practice, we perform this projection by constructing its singular value decomposition, which for a non-square matrix takes the form $\tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger \boldsymbol{W}_- = \boldsymbol{U} \boldsymbol{D} \boldsymbol{V}^\dagger$. Here, $\boldsymbol{D}$ is a diagonal matrix with the same dimension as $\tilde{\boldsymbol{D}}_+$, $\boldsymbol{U}$ is a unitary matrix of the same dimension, and $\boldsymbol{V}$ is a column orthogonal rectangular matrix. There are at most two non-zero singular values (elements of the diagonal matrix $\boldsymbol{D}$) per m, and the corresponding left singular vectors (columns of $\boldsymbol{U}$) form an orthonormal basis for the range of $\tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger \boldsymbol{W}_-$. We can project out of this range by defining $\tilde{\boldsymbol{U}}$ as the matrix obtained by removing the columns of $\boldsymbol{U}$ where the corresponding singular value is non-zero. Thus, choosing the window functions

$$\boldsymbol{W}^* = \tilde{\boldsymbol{U}}^\dagger \tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger, \tag{35}$$

we guarantee separation of the electric and magnetic polarization for azimuthally-symmetric patches. Our separated polarization observables become

$$\mathbf{B}_W = \boldsymbol{W}^* \tilde{\mathbf{B}} \equiv \tilde{\boldsymbol{U}}^\dagger \tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger \tilde{\mathbf{B}} \approx \tilde{\boldsymbol{U}}^\dagger \tilde{\boldsymbol{D}}_+^{1/2} \tilde{\boldsymbol{U}}_+^\dagger \mathbf{B}, \tag{36}$$

$$\mathbf{E}_W = \boldsymbol{W}^* \tilde{\mathbf{E}} \equiv \tilde{\boldsymbol{U}}^\dagger \tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger \tilde{\mathbf{E}} \approx \tilde{\boldsymbol{U}}^\dagger \tilde{\boldsymbol{D}}_+^{1/2} \tilde{\boldsymbol{U}}_+^\dagger \mathbf{E}. \tag{37}$$

For azimuthal patches the separation is exact; the approximation sign arises only from our use of $\boldsymbol{W}_+ \approx \tilde{\boldsymbol{U}}_+ \tilde{\boldsymbol{D}}_+ \tilde{\boldsymbol{U}}_+^\dagger$ in simplifying the matrix that premultiplies $\mathbf{E}$ and $\mathbf{B}$ in Eqs. (36) and (37).

For observations over non-azimuthal patches, one can either apply the exact separation over the largest inscribed azimuthal region or attempt to apply the method outlined above to the entire patch. In the latter case the structure of $\boldsymbol{W}_-$ is less clear, but we can still expect a significant number of its eigenvalues to be very small; the associated eigenvectors correspond to window functions that satisfy $W \approx 0 \approx \eth W$ on the boundary. The number of independent window functions that achieve separation can be estimated as the number of pixels of linear size $\sim \pi / l_{\rm max}$ contained in the patch (roughly the number of modes that survive the diagonalization of $\boldsymbol{W}_+$) minus twice the number of pixels on the boundary (roughly the number of constraints in setting W and $\eth W$ to zero on the boundary). Such window functions will only give an approximate separation of electric and magnetic polarization. In practice, the accuracy of the separation, and the number of independent window functions constructed, will depend on the choice of threshold for retaining the singular values of $\tilde{\boldsymbol{D}}_+^{-1/2} \tilde{\boldsymbol{U}}_+^\dagger \boldsymbol{W}_-$.

If we are only interested in constructing variables that depend on the magnetic polarization, the maximum multipole $l_{\rm max}$ in the window functions can be chosen rather small (of the order of a few hundred). The relation $\tilde{\mathbf{B}} = \boldsymbol{W}_+ \mathbf{B} - i \boldsymbol{W}_- \mathbf{E}$ will only hold for square $\boldsymbol{W}_-$ if $l_{\rm max}$ is chosen to include all the significant power in the electric polarization, so for smaller $l_{\rm max}$ we cannot assume that $\boldsymbol{W}_-$ is square. However, for azimuthal patches, the range of each submatrix of $\boldsymbol{W}_-$ is still guaranteed to be of dimension two or less, irrespective of its shape, and so the exact separation can proceed with $\boldsymbol{W}_-$ treated as square. For non-azimuthal patches it would be prudent to monitor the effect of varying the number of columns in $\boldsymbol{W}_-$ on the range of this matrix.

A. Noise

It is straightforward to project the errors on the Stokes' parameters to find the noise in the separated variables $\mathbf{E}_W$ and $\mathbf{B}_W$. In this section we consider the simple case of maps with idealized noise properties.

We assume that the noise correlation between pixels is negligible, and that the noise on the Stokes' parameters ΔQ and ΔU is un-correlated. The neglect of noise correlations between pixels amounts to assuming white noise in the time stream of the measurement. In principle, correlations between Stokes' parameters can be eliminated with a careful choice of polarimeter directions in the experiment [19, 20]. With these assumptions, we have

$$\langle \Delta Q(\Omega)\Delta U(\Omega')\rangle = 0, \tag{38}$$

and the noise correlation is given by

$$\langle \Delta Q(\Omega)\Delta Q(\Omega')\rangle = \sigma_Q^2\delta(\Omega-\Omega'), \qquad \langle \Delta U(\Omega)\Delta U(\Omega')\rangle = \sigma_U^2\delta(\Omega-\Omega'). \tag{39}$$

It follows that the noise $\Delta\tilde{\mathbf{B}}$ on the pseudo-multipoles $\tilde{\mathbf{B}}$ has correlations

$$\langle \Delta\tilde{\mathbf{B}}\Delta\tilde{\mathbf{B}}^\dagger\rangle = \tfrac{1}{2}\int_S \mathrm{d}S\,(\sigma_Q^2+\sigma_U^2)\tfrac{1}{2}[{}_{-2}\mathbf{Y}^*({}_{-2}\mathbf{Y}^*)^\dagger + {}_2\mathbf{Y}^*({}_2\mathbf{Y}^*)^\dagger] - \tfrac{1}{2}\int_S \mathrm{d}S\,(\sigma_Q^2-\sigma_U^2)\tfrac{1}{2}[{}_{-2}\mathbf{Y}^*({}_2\mathbf{Y}^*)^\dagger + {}_2\mathbf{Y}^*({}_{-2}\mathbf{Y}^*)^\dagger], \tag{40}$$

with a similar result for the noise $\Delta\tilde{\mathbf{E}}$ on $\tilde{\mathbf{E}}$. Here, vectors of spin-weight s functions ${}_s\mathbf{Y}$ have components ${}_sY_{lm}$. The polarimeter arrangements that give uncorrelated errors between Stokes' parameters also ensure that $\sigma_U^2=\sigma_Q^2\equiv\sigma_N^2$ so that the last integral in Eq. (40) is zero. Here we concentrate on the simple case where σ_N^2 is isotropic in which case

$$\langle \Delta\tilde{\mathbf{B}}\Delta\tilde{\mathbf{B}}^\dagger\rangle = \langle \Delta\tilde{\mathbf{E}}\Delta\tilde{\mathbf{E}}^\dagger\rangle = \sigma_N^2\boldsymbol{W}_+. \tag{41}$$

The covariance of the noise $\Delta\mathbf{B}_W$ on $\mathbf{B}_W$ is therefore given by

$$\boldsymbol{N} \equiv \langle \Delta\mathbf{B}_W\Delta\mathbf{B}_W^\dagger\rangle = \boldsymbol{W}^*\langle\Delta\tilde{\mathbf{B}}\Delta\tilde{\mathbf{B}}^\dagger\rangle(\boldsymbol{W}^*)^\dagger \approx \sigma_N^2\tilde{\boldsymbol{U}}^\dagger\tilde{\boldsymbol{D}}_+^{-1/2}\tilde{\boldsymbol{U}}_+^\dagger\tilde{\boldsymbol{U}}_+\tilde{\boldsymbol{D}}_+\tilde{\boldsymbol{U}}_+^\dagger\tilde{\boldsymbol{U}}_+\tilde{\boldsymbol{D}}_+^{-1/2}\tilde{\boldsymbol{U}} = \sigma_N^2\boldsymbol{I}, \tag{42}$$

and hence the noise is diagonal (and isotropic); similarly $\langle\Delta\mathbf{E}_W\Delta\mathbf{E}_W^\dagger\rangle=\sigma_N^2\boldsymbol{I}$. What is more, the noise on $\mathbf{E}_W$ and $\mathbf{B}_W$ are uncorrelated for isotropic noise since

$$\langle \Delta\tilde{\mathbf{E}}\Delta\tilde{\mathbf{B}}^\dagger\rangle = i\sigma_N^2\boldsymbol{W}_-, \tag{43}$$

and hence

$$\langle \Delta\mathbf{E}_W\Delta\mathbf{B}_W^\dagger\rangle = i\sigma_N^2\tilde{\boldsymbol{U}}^\dagger\tilde{\boldsymbol{D}}_+^{-1/2}\tilde{\boldsymbol{U}}_+^\dagger\boldsymbol{W}_-\tilde{\boldsymbol{U}}_+\tilde{\boldsymbol{D}}_+^{-1/2}\tilde{\boldsymbol{U}} = 0 \tag{44}$$

as $\tilde{\boldsymbol{U}}^\dagger$ annihilates $\tilde{\boldsymbol{D}}_+^{-1/2}\tilde{\boldsymbol{U}}_+^\dagger\boldsymbol{W}_-$. For isotropic noise our polarization variables therefore have the same desirable diagonal properties as the scalar diagonalized harmonic coefficients used in the analysis of the cut-sky CMB temperature [10, 11].

In the presence of a symmetric beam, white noise in the time stream of the experiment projects to white (though generally non-isotropic) noise on the beam-convolved polarization field. In multipole space, the convolved fields have electric and magnetic multipoles that are related to the unconvolved E_{lm} and B_{lm} by spin-weight 2 beam window functions ${}_2W_l$ [21, 22]. For $l_{\max}$ of a few hundred, appropriate for probing magnetic polarization, the beam window functions will be negligible for experiments with resolution much better than one degree. For lower resolution experiments, or higher $l_{\max}$, the effect of the beam window function should be included in the theoretical (signal) covariance of the variables $\mathbf{E}_W$ and $\mathbf{B}_W$ (see below).

In general, non-uniform coverage of the sky will lead to variations in σ_N^2. In this case it is still possible to define harmonic variables $\mathbf{B}_W^{\rm diag}=\boldsymbol{N}^{-1/2}\mathbf{B}_W$ that have isotropic noise. (Here $\boldsymbol{N}^{-1/2}=\boldsymbol{U}_N\boldsymbol{D}_N^{-1/2}\boldsymbol{U}_N^\dagger$ where $\boldsymbol{N}=\boldsymbol{U}_N\boldsymbol{D}_N\boldsymbol{U}_N^\dagger$, $\boldsymbol{U}_N$ is unitary and $\boldsymbol{D}_N$ is diagonal.) For azimuthal patches with general (everywhere finite) noise patterns we can still construct the magnetic-only variables for each value of m but the noise will now couple variables with different m (unlike the signal). While presenting no fundamental obstacles, this m-mode coupling does increase the computational overhead considerably.

Sky patches of general shape are equivalent to azimuthal patches with regions of infinite noise, and the general case of a non-azimuthally symmetric survey region with arbitrary noise can be treated by re-defining the pseudo-harmonics $\tilde{\mathbf{E}}$, $\tilde{\mathbf{B}}$ and coupling matrices $\boldsymbol{W}_+$, $\boldsymbol{W}_-$ with a factor of $1/\sigma_N^2$ inside the integral. All the above results then follow with the new definitions, though computing the matrices and manipulating them may become computationally challenging. The assumption of isotropic noise is therefore not fundamental to our analysis, and the following results could be generalized for more realistic situations.

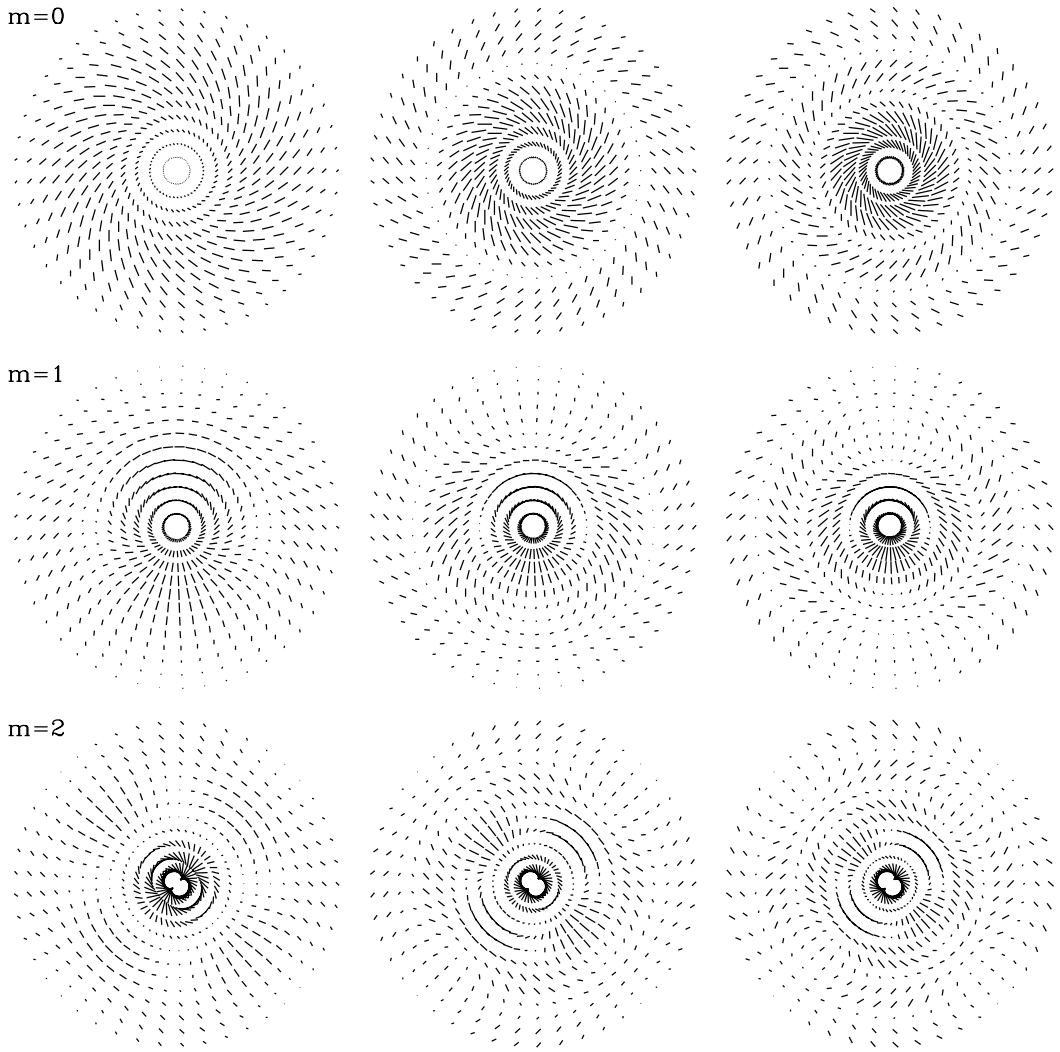

FIG. 3: The real space window functions for an azimuthally symmetric sky patch with $\theta < 10^\circ$. They are evaluated in the frame where the signal is diagonal, so the leftmost window produces the largest signal for that m. The signal to noise falls off as a function of m as shown in Fig. 4. For $m > 0$ the window functions correspond to the real part of $\mathbf{B}_W^{(R)}$; the imaginary part is a rotated version of the same window. The length of the lines shows the sampling weight at that point, and the orientation of the lines shows which polarization orientation gives maximal contribution.

B. Real space window functions

The expected magnetic signal correlation matrix is given by

$$\boldsymbol{S} \equiv \langle \mathbf{B}_W \mathbf{B}_W^\dagger \rangle = \tilde{\boldsymbol{U}}^\dagger \tilde{\boldsymbol{D}}_+^{1/2} \tilde{\boldsymbol{U}}_+^\dagger \boldsymbol{C}^{BB} \tilde{\boldsymbol{U}}_+ \tilde{\boldsymbol{D}}_+^{1/2} \tilde{\boldsymbol{U}}, \tag{45}$$

where the diagonal magnetic power spectrum matrix is given by $[\boldsymbol{C}^{BB}]_{(lm)(l'm')} = \delta_{mm'}\delta_{ll'}C_l^{BB}$. Since the noise correlation is proportional to the identity matrix for isotropic noise we can perform any rotation $\mathbf{B}_W \to \mathbf{B}_W^{(R)} \equiv \boldsymbol{R}\mathbf{B}_W$, where $\boldsymbol{R}$ is unitary, and still have a set of variables with uncorrelated errors. The rotated variables are derived from window functions $\boldsymbol{R}^*\boldsymbol{W}$. For a particular theoretical model we can rotate to the frame where the signal matrix is diagonal. The rotated $\mathbf{B}_W^{(R)}$ will then be fully statistically independent. In Fig. 3 we plot the window functions for the $\mathbf{B}_W^{(R)}$ which give the largest contributions to the signal for a typical flat ΛCDM model with a scale invariant tensor initial power spectrum and no reionization. The window functions are plotted as line segments of length $\sqrt{Q_W^2 + U_W^2}$

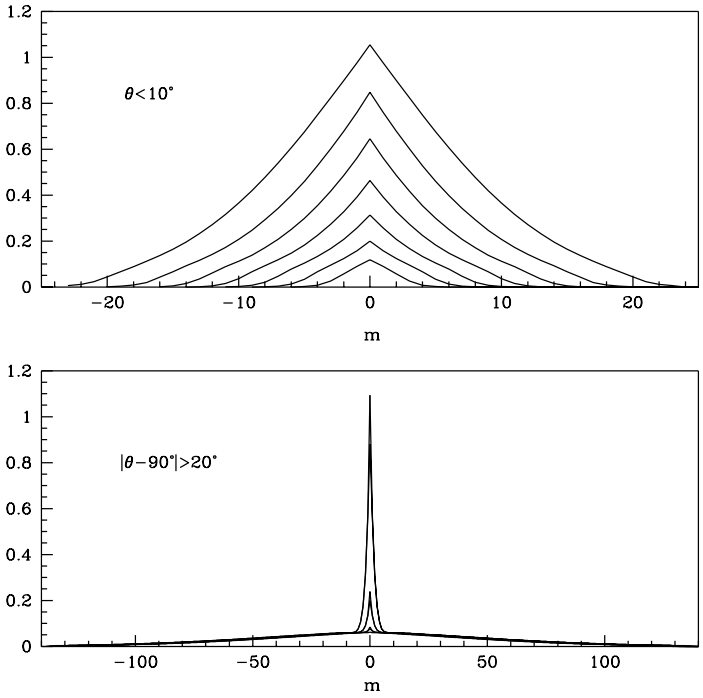

FIG. 4: The eigenvalues of $\boldsymbol{N}^{-1/2}\boldsymbol{S}\boldsymbol{N}^{-1/2}$ at a given m with the tensor amplitude chosen to give a detection at 99% confidence with probability 0.5 (see Sec. IV). The noise is assumed isotropic and the model has reionization at $z = 6.5$. For the small sky patch most of the signal is in the lowest few eigenmodes of each m, but for larger patches a greater number of eigenmodes are required to encompass all the signal (in the bottom plot there are lots of contributions on top of each other along the bottom line). For the large patch there are a small number of high signal to noise modes due to the extra large scale reionization signal.

at angle $\tan^{-1}(U_W/Q_W)/2$ to the θ direction where the real quantities Q_W and U_W are defined in terms of the real part $\Re W^{(R)}$ of the (rotated) scalar window function $W^{(R)}$ as

$$Q_W + iU_W \equiv -i\,\overline{\eth}\,\overline{\eth}\,\Re W^{(R)}. \tag{46}$$

This definition ensures that

$$\Re\mathbf{B}_W^{(R)} = \int_S \mathrm{d}S\,(Q_W Q + U_W U)\,. \tag{47}$$

For the imaginary part of $\mathbf{B}_W^{(R)}$, Q_W and U_W should be defined as $Q_W + iU_W \equiv -i\,\overline{\eth}\,\overline{\eth}\,\Im W^{(R)*}$. For the case of azimuthal patches, as considered in Fig. 3 where the windows are constructed for each m, the imaginary part would produce a plot that is rigidly rotated by $-\pi/(2m)$ ($m \neq 0$) about the centre. Plotting the window functions in this form is useful since the length of the line segment gives the sampling weight assigned to that point, and the orientation gives the direction of the linear polarization that contributes at each point. We could repeat the exercise for the $\mathbf{E}_W$ in which case for the real part we would define $Q_W + iU_W \equiv \overline{\eth}\,\overline{\eth}\,\Re W^{(R)}$.

In Fig. 4 we show the signal to noise in the magnetic variables for two azimuthal patches. As the patch size increases the signal in the modes with large m also increases, reflecting the fact that for small patches the diagonalization of $\boldsymbol{W}_+$ removes a greater relative fraction of the modes at each m as m increases. For small patches of the sky most of the signal at each m is compressed into a small number of modes, whereas for larger patches the signal is distributed more uniformly. For cosmological models with reionization the signal for large patches is distributed less uniformly, with a small number of modes giving big contributions due to the greater large scale power.

IV. MEASURING THE MAGNETIC SIGNAL

We are now in a position to use the magnetic observable $\mathbf{B}_W$ to constrain the magnetic signal without having to worry about contamination with the much larger electric signal. The simplest thing to do would be to test the null

hypothesis that the magnetic signal is due entirely to noise (this hypothesis is unlikely to be ruled out pre-Planck). If the signal were not consistent with noise it could indicate various things: the presence of CMB magnetic polarization, the presence of polarized foregrounds, that haven't been removed successfully, systematic leakage into the magnetic mode in the analysis (e.g. due to unaccounted for pointing errors, or pixelization effects), or Q-U leakage in the observation (e.g. due to unaccounted for cross-polarization in the instrument optics).

Magnetic polarization can originate from tensor modes, but also by weak lensing of the scalar electric polarization [23]. The lensing signal should be dominant on small scales, and the magnetic variables could certainly be used to observe this signal. Of more interest here is the larger scale contribution from tensor modes. In order to identify this component we shall have to model the lensing contribution, which becomes increasingly important as one tries to observe smaller tensor contributions [7, 8]. In the first three of the following subsections we assume that the magnetic signal is generated purely from the tensor modes, then in subsection IV D we show how our results can be adapted to account for the lensing signal.

A. Is it just noise?

If the noise and signal are Gaussian the $\mathbf{B}_W$ will be Gaussian and the simplest thing to do is a χ^2 test by computing $\chi^2 = \mathbf{B}_W^\dagger \boldsymbol{N}^{-1}\mathbf{B}_W$ (for isotropic noise this is just $\chi^2 = \mathbf{B}_W^\dagger \mathbf{B}_W/\sigma_N^2$). Whilst the CMB magnetic polarization signal is from tensor modes is expected to be Gaussian, the lensing signal and any spurious or unexpected signal may not be. One may therefore also wish to do a more sophisticated set of statistical tests at this point.

Assuming that the signal is as expected — any B signal present is Gaussian and would have a power spectrum as predicted for a near scale-invariant tensor initial power spectrum — one can account for the expected form of the power spectrum and thereby increase the chance of a detection. We assume that the main parameters of the universe are well determined by the time magnetic polarization comes to be observed, so the shape of the magnetic polarization power spectrum is known to reasonable approximation (the only significant freedom arising from the shape of the primordial tensor power spectrum). We compute the expected signal correlation $\boldsymbol{S}$ for some particular tensor amplitude and say that the real signal is $r\boldsymbol{S}$. Assuming Gaussian signal and noise the likelihood in this case is then given by

$$L(\mathbf{B}_W|r) \propto \frac{\exp[-\frac{1}{2}\mathbf{B}_W^\dagger(\boldsymbol{N}+r\boldsymbol{S})^{-1}\mathbf{B}_W]}{|\boldsymbol{N}+r\boldsymbol{S}|^{1/2}}. \tag{48}$$

The likelihood distribution can be computed numerically from the $\mathbf{B}_W$ observed, and gives the posterior probability distribution on the value of r after multiplying by the prior $f(r)$.

The magnetic signal is expected to be weak, and the detailed statistics for analysing such a signal are given in Appendix D. There we show that

$$\nu' \equiv \frac{\mathbf{B}_W^\dagger \boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\mathbf{B}_W - \mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S})}{\sqrt{4\mathbf{B}_W^\dagger \boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\mathbf{B}_W - 2\mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S})}}. \tag{49}$$

gives a measure of the number of 'sigmas' of the detection — the number of standard deviations of the maximum likelihood $\hat{r}$ from pure noise ($r = 0$) assuming low signal to noise. We use this as a test statistic in Monte-Carlo simulations to compute detection probabilities at a given significance. We have checked at isolated points that using optimal statistics gains very little except for very small sky patches (where there are only a small number of magnetic modes, each of which must have fairly high signal to noise in order to get a detection).

Using the $\mathbf{B}_W$ variables is clearly not optimal as we have thrown away some well determined linear combinations of E and B. However in the idealized situation considered here they should provide a robust way for testing for magnetic polarization. The number of modes thrown away is in any case quite small — not more than two per m mode for azimuthal patches. We quantify this information loss further in Sec. V.

B. Detection by Planck?

Of the current funded experiments, only Planck is likely to detect magnetic polarization if the levels are as predicted by standard cosmological models. As a toy model we consider the 143 and 217 GHz polarized channels of the Planck High Frequency Instrument. We approximate the noise as isotropic and ignore the variation of beam width (7.1 and 5.0 arcmin full width at half maximum respectively) between these channels. Combining maps from these two channels with inverse variance weighting, we find $\sigma_N \approx 6\times10^{-3}\mu\mathrm{K/K}$, where Q and U are expressed as dimensionless

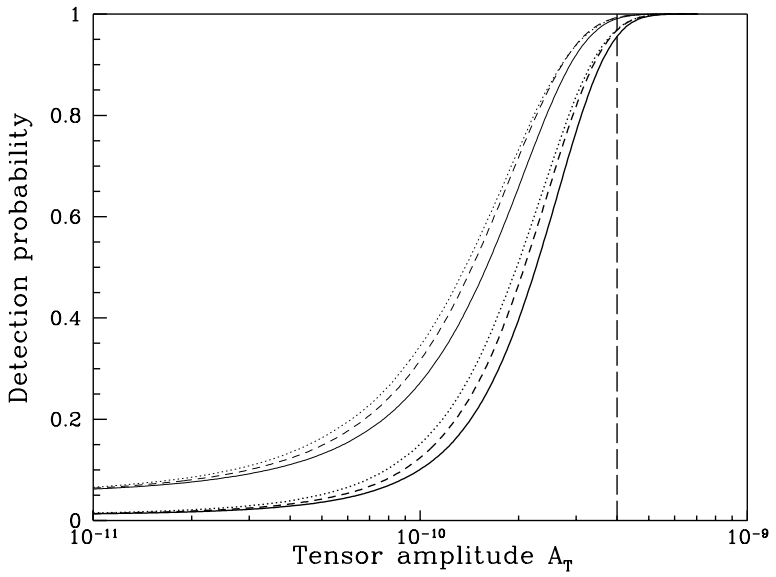

FIG. 5: The probability of detecting magnetic polarization at 99 per cent (thick lines) and 95 per cent (thin lines) confidence as a function of the tensor initial power spectrum amplitude A_T for the model of Planck observations described in the text. The dotted line is for a model with reionization at $z = 6.5$, using the unprojected variables and treating the electric contamination as part of the noise (see Sec. V). The other curves are using the projected magnetic variables for models with no reionization (solid) and reionization at $z = 6.5$ (dashed). The vertical dashed line show the tensor amplitude that would contribute about 1/10 of the COBE signal.

thermodynamic equivalent temperatures in units of the CMB temperature. We apply an azimuthally-symmetric galactic cut of 20 degrees either side of the equator. The expected magnetic polarization power spectrum peaks at $l \sim 100$, and there is therefore no need to consider high resolutions so we can use $l_{\rm max} = 250$ without significant loss of power. In Fig. 5 we show the probability of obtaining a detection with Planck as a function of the true underlying scale-invariant tensor power spectrum amplitude A_T (defined as in Ref. [24]) assuming a standard flat ΛCDM model.

A tensor amplitude of $A_T \approx 4 \times 10^{-10}$ would contribute about 1/10 of the large scale temperature C_l detected by COBE, and is likely to be detected by Planck if our model is at all realistic. This corresponds to being able to detect the signal from inflationary models with energy scale at horizon crossing $V^{1/4} \gtrsim 2 \times 10^{16}\,{\rm GeV}$. Such models include the simple ϕ^n potentials, with $n \geq 2$.

C. Survey size

For a given detector sensitivity the magnitude of the signal that can be detected depends on the size of the sky patch that is observed. The signal to noise in each observable increases in proportion to the observation time per unit area. The noise covariance is proportional to σ_N^2 which varies in proportion to the observed area for a given survey duration. For large areas the number of observables varies approximately in proportion to the area, which would make the number of 'sigmas' of a chi-squared detection scale with the square root of the area. Combining these two effects, the expected detection is therefore proportional to one over the square root of the area, and is larger if a fixed survey time is spent observing a smaller area. However for smaller areas the signal to noise on each observable becomes larger, and the number of variables decreases. With fewer variables the probability of obtaining no detection increases significantly. This is just the fact that if you observe a small patch of sky you have a larger chance of being unlucky and having a patch which has a small magnetic polarization signal everywhere. Also the existence of the boundary becomes increasingly important for small patches and a larger fraction of the information is lost in order to obtain clean separation of the magnetic observables.

The question of 'optimal' survey size is somewhat delicate, as it depends on the probability distribution for the detection significance that one thinks is optimal. In Fig. 6 we plot the probability of detecting various tensor amplitudes at 95 per cent and 99 per cent confidence for different survey sizes. In Fig. 7 we show the minimum gravitational wave (tensor) amplitude that might be detected at 99 per cent confidence as a function of the radius of the survey size. It is clear that radii in the range 5°–9° are optimal, though one cannot be more precise without defining more

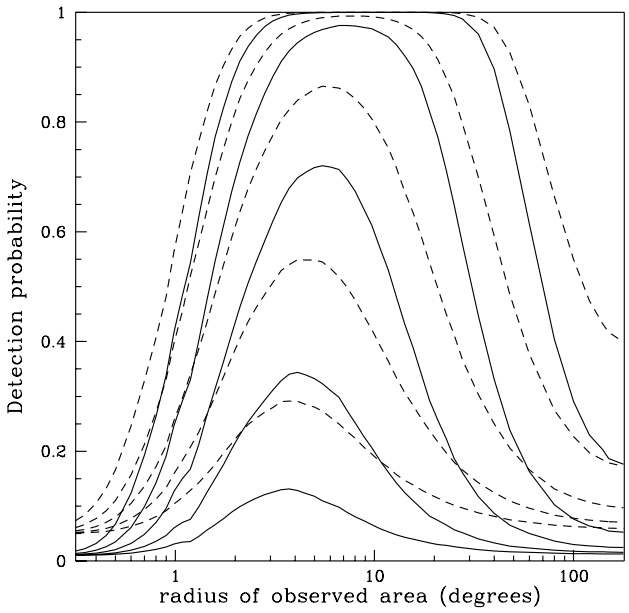

FIG. 6: The probability of being able to rule out the null hypothesis at 95 per cent (dashed lines) or 99 per cent (solid lines) confidence for scale invariant tensor amplitudes (bottom to top) of $A_T = \{1, 2, 4, 8, 16\} \times 10^{-12}$, for a survey with detector sensitivity $s = 10\,\mu\mathrm{K}\sqrt{\mathrm{sec}}$ that runs for one year and maps a circular patch of sky of a given radius assuming uniform noise, no reionization, and no lensing. The probabilities were computed by Monte-Carlo simulation.

specifically the aims of the observation. A radius of about 7° would be a good compromise between being able to place good upper limits if there was no detection (which favours radii closer to 9°) and having a better chance of detecting small amplitudes (which favours smaller radii).

The solid curves in Fig. 7 fully take account of the need to separate the magnetic signal from the (much larger) electric signal. By way of comparison, the dashed curves show the minimum detectable amplitude obtainable if one could do perfect lossless separation, which is clearly impossible on an incomplete sky (see Sec. V). With lossless separation, the best upper bounds are obtained for smaller patches since the size of the boundary is no longer important. The dashed curves in Fig. 7 can be compared with those given in Ref. [25] where perfect separation was assumed, the effects of finite sky coverage were treated only approximately, and a less rigorous approach to hypothesis testing was employed. Ref. [26] gives an improved analysis along the lines of Ref. [25], and also performs calculations properly taking account of the mixing of electric and magnetic polarization through a (brute-force) Fisher analysis in pixel space.

D. Lensing

Unlike most of the foreground signals that might contaminate the observation, the magnetic signal from the lensing of scalar electric polarization has the same frequency spectrum as the primordial magnetic signal and so cannot be removed easily by use of multi-frequency observations. In order to isolate the tensor contribution to the magnetic signal we can incorporate knowledge of the expected lensing power spectrum [23] into the null-hypothesis covariance matrix $\boldsymbol{N}$ (we neglect the non-Gaussianity of the lensed polarization). For the multipoles $l \lesssim 250$ of interest for the tensor signal the lensing signal is approximately white, with $C_{\mathrm{lens}}^{BB} \approx 4.4 \times 10^{-6}\mu\mathrm{K}^2$ if the COBE signal is entirely generated by scalar modes. For large patches of sky, where the matrix $\tilde{\boldsymbol{D}}_+$ is nearly proportional to the identity matrix, the lensing signal contributes like an additional isotropic noise with $\sigma_{N,\mathrm{lens}}^2 \approx C_{\mathrm{lens}}^{BB}$. We have checked this approximation by computing the following results exactly in particular cases, with agreement to within a few percent for patch sizes with $10° \lesssim \theta \lesssim 80°$.

The effect of the lensing is therefore simply to increase the effective noise by a constant amount. For the Planck satellite the effect is small, reducing the A_T that could be observed by about 1.5 per cent. However for the smaller surveys with better sensitivity, considered in Figs. 6 and 7, the effect is much more important. For a one year survey of radius θ with sensitivity $s = 10\,\mu\mathrm{K}\sqrt{\mathrm{sec}}$ the noise variance is given by $f(\theta)\sigma_0^2$ where $f(\theta) = (1 - \cos\theta)/2$ is the

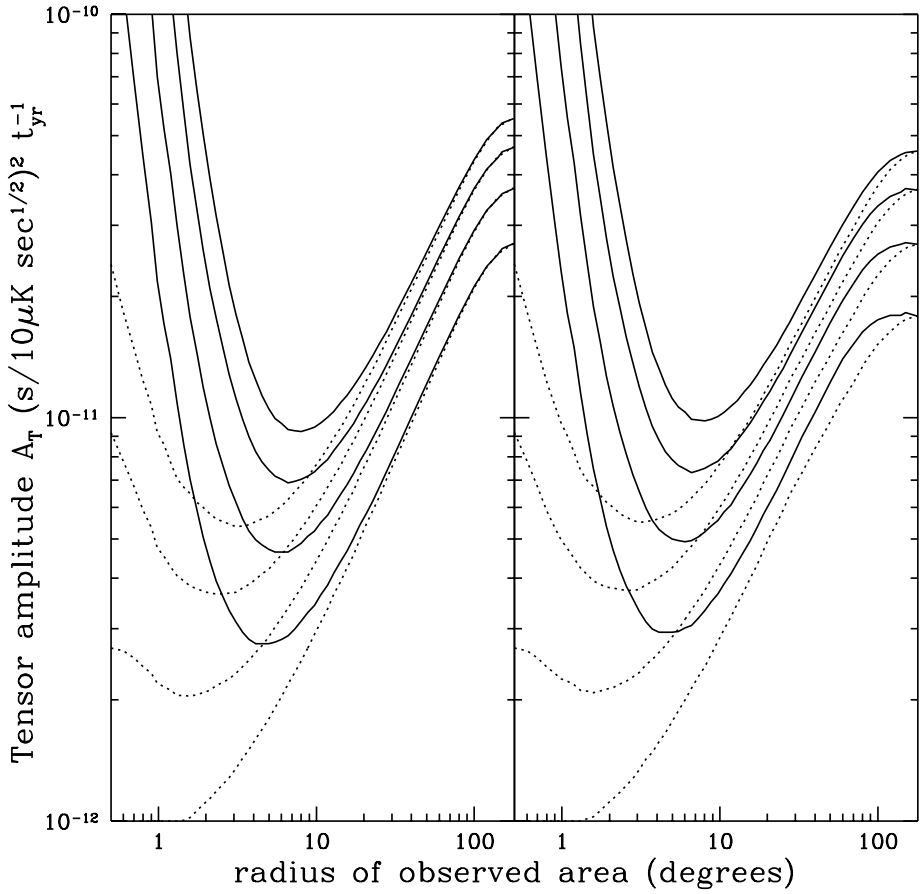

FIG. 7: The smallest gravitational wave amplitude A_T (defined as in Ref. [24]) that could be detected at 99 per cent confidence with probability (bottom to top) of {0.5,0.8,0.95,0.99} by an experiment with detector sensitivity $s = 10\,\mu\mathrm{K}\sqrt{\mathrm{sec}}$ that runs for one year and maps a circular patch of sky of a given radius assuming uniform noise, no lensing, with reionization at $z = 6.5$ (right) and without reionization (left). The result scales with the square of the detector sensitivity and inversely with the duration of the experiment. The dotted lines show the equivalent result if one could perform perfect lossless separation of the magnetic modes.

fraction of the sky which is observed and $\sigma_0^2 \approx 4 \times 10^{-5}\mu\mathrm{K}^2$. This noise gives the tensor amplitudes $A_T(\theta)$ plotted in Fig. 7. Incorporating the lensing effect the actual tensor amplitude one could detect in an experiment with sensitivity s and duration T is

$$A_T(\theta)' = A_T(\theta)\frac{4\pi f(\theta)s^2/T + \sigma_{N,\mathrm{lens}}^2}{f(\theta)\sigma_0^2}, \tag{50}$$

where $A_T(\theta)$ is the amplitude for a one year mission with $s = 10\,\mu\mathrm{K}\sqrt{\mathrm{sec}}$ and lensing ignored (i.e. the amplitude plotted in Fig. 7). This allows our previous results to be modified for inclusion of the lensing signal. We have plotted the modified results for various survey sensitivities in Fig. 8. The optimal survey size now depends on the sensitivity—as sensitivity improves the lensing signal becomes more important and one needs to survey larger scales in order to accurately measure the difference in variance expected with the tensor signal. For large patch sizes the tensor amplitude A_T that can be detected in the absence of lensing is proportional to $\sqrt{f(\theta)}$. Allowing for lensing, there is an optimal survey size at $4\pi s^2 f(\theta)/T = \sigma_{N,\mathrm{lens}}^2$ [if there is a solution with $f(\theta) < 1$], in other words when the variance of the instrument noise is equal to the lensing signal. There is a lower limit of $A_T \approx 4 \times 10^{-12}$ that can be measured even with perfect sensitivity, when the tensor contribution cannot be distinguished from random sampling variations in the lensing signal distribution. This corresponds to an inflation model with energy scale $V^{1/4} \approx 7 \times 10^{15}\,\mathrm{GeV}$, in broad agreement with Ref. [7] for a three sigma detection. This situation could only be improved if one could find ways to obtain information about the particular realization of the lensed signal.

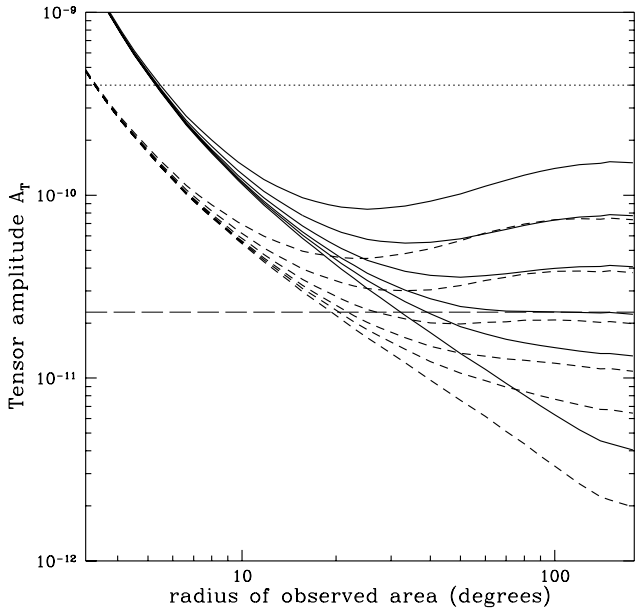

FIG. 8: The smallest gravitational wave amplitude A_T that could be detected at 99 per cent confidence with probability 0.95 (solid lines) and 0.5 (dashed lines) by an experiment with detector sensitivity (bottom to top) $s^2 = \{0, 25, 50, 100, 200, 400\}\ \mu\mathrm{K}^2\mathrm{sec}$ that runs for year and maps a circular patch of sky of a given radius assuming uniform noise, a white lensing power spectrum, and a ΛCDM cosmology that reionizes at $z = 6.5$. The horizontal dotted line shows the amplitude which contributes about 1/10 of the COBE signal. The long dashed line shows very roughly the dust foreground at 143 GHz.

We have assumed that component separation and source subtraction can be performed exactly so that the observed signal is only lensed CMB. Polarized thermal dust emission is expected to generate a significant magnetic signal [27] at roughly the level shown by the line in Fig. 8 at 143 GHz. Separation of this signal from the CMB signal should be possible with multi-frequency observations, and it should then not have a significant effect on our results.

V. LOSSLESS METHODS

We now compare the above analysis with truly lossless methods. Lossless, likelihood analysis for CMB polarization in pixel space has been considered recently in Ref. [26]. In this section we consider lossless and nearly lossless methods in harmonic space.

A simple way to incorporate most of the magnetic signal for constraining the tensor modes is to use the unprojected variables $\tilde{\boldsymbol{D}}_+^{-1/2}\tilde{\boldsymbol{U}}_+^\dagger\tilde{\mathbf{B}}$, where $\tilde{\boldsymbol{D}}_+$ and $\tilde{\boldsymbol{U}}_+$ were defined in Sec. III. The null-hypothesis covariance matrix $\boldsymbol{N}$ can be computed including the expected signal from the electric polarization, and the analysis can be performed as before. This is marginally superior to using the projected variables if the tensor amplitude is quite high, as shown in Fig. 5 for the Planck mission. For smaller tensor amplitudes the entangled linear combinations of E and B modes are dominated by the electric component and performing the projection looses very little. Using the projection gives one clean separation, and there is no need to know the electric polarization power spectrum. By identifying variables that depend only on the electric and magnetic signals at the level of the map we also do not need to assume Gaussianity, so we could for example perform Gaussianity tests on the two physically distinct polarization types independently.

We now consider the full joint analysis of the electric and magnetic polarization, with the pseudo-multipoles $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{B}}$ forming our fundamental data vector

$$\begin{pmatrix}\tilde{\mathbf{E}}\\ \tilde{\mathbf{B}}\end{pmatrix} = \begin{pmatrix}\boldsymbol{W}_+ & i\boldsymbol{W}_-\\ -i\boldsymbol{W}_- & \boldsymbol{W}_+\end{pmatrix}\begin{pmatrix}\mathbf{E}\\ \mathbf{B}\end{pmatrix}. \tag{51}$$

Since we are no longer worrying about E-B separation so we can equally well use the block-diagonal frame where

$$\tilde{\mathbf{A}} \equiv \begin{pmatrix} \tilde{\mathbf{E}} + i\tilde{\mathbf{B}} \\ \tilde{\mathbf{E}} - i\tilde{\mathbf{B}} \end{pmatrix} = \begin{pmatrix} {}_2\boldsymbol{W} & \mathbf{0} \\ \mathbf{0} & {}_{-2}\boldsymbol{W} \end{pmatrix} \begin{pmatrix} \mathbf{E} + i\mathbf{B} \\ \mathbf{E} - i\mathbf{B} \end{pmatrix} \equiv \begin{pmatrix} {}_2\boldsymbol{W} & \mathbf{0} \\ \mathbf{0} & {}_{-2}\boldsymbol{W} \end{pmatrix} \mathbf{A}. \tag{52}$$

Performing a singular value decomposition of the block-diagonal matrix $\mathrm{diag}({}_2\boldsymbol{W}, {}_{-2}\boldsymbol{W}) = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^\dagger$ (where the matrices $\boldsymbol{U}$ and $\boldsymbol{D}$ should not be confused with those defined in Sec. III), we can identify the well determined linear combinations as before

$$\mathbf{A}' \equiv \tilde{\boldsymbol{D}}^{-1/2}\tilde{\boldsymbol{U}}^\dagger\tilde{\mathbf{A}} \approx \tilde{\boldsymbol{D}}^{1/2}\tilde{\boldsymbol{U}}^\dagger\mathbf{A}. \tag{53}$$

This diagonalization is equivalent to defining harmonic coefficients $\mathbf{A}'$ with respect to a complete set of spin two harmonics which are orthonormal over the patch of sky, in the same way as for the spin zero cut-sky temperature analysis [10, 11]. As before, this construction ensures that for isotropic noise the noise correlation is diagonal

$$\langle \Delta\tilde{\mathbf{A}}\Delta\tilde{\mathbf{A}}^\dagger \rangle = 2\sigma_N^2 \mathrm{diag}({}_2\boldsymbol{W}, {}_{-2}\boldsymbol{W}) \quad \Longrightarrow \quad \boldsymbol{N} \equiv \langle \Delta\mathbf{A}'\Delta\mathbf{A}'^\dagger \rangle = 2\sigma_N^2\boldsymbol{I}. \tag{54}$$

The signal correlation is given by

$$\boldsymbol{S} = \langle \mathbf{A}'\mathbf{A}'^\dagger \rangle = \tilde{\boldsymbol{D}}^{1/2}\tilde{\boldsymbol{U}}^\dagger \begin{pmatrix} \boldsymbol{C}^{EE} + \boldsymbol{C}^{BB} & \boldsymbol{C}^{EE} - \boldsymbol{C}^{BB} \\ \boldsymbol{C}^{EE} - \boldsymbol{C}^{BB} & \boldsymbol{C}^{EE} + \boldsymbol{C}^{BB} \end{pmatrix} \tilde{\boldsymbol{U}}\tilde{\boldsymbol{D}}^{1/2}, \tag{55}$$

where the $\boldsymbol{C}^{EE}$ and $\boldsymbol{C}^{BB}$ are the diagonal electric and magnetic power spectrum matrices respectively, and we have assumed that $\boldsymbol{C}^{EB} = 0$. If the noise and signal are Gaussian we can proceed to do a likelihood analysis for the power spectra using

$$L(\mathbf{A}'|\boldsymbol{C}^{EE}, \boldsymbol{C}^{BB}) \propto \frac{\exp[-\frac{1}{2}\mathbf{A}'^\dagger(\boldsymbol{N} + \boldsymbol{S})^{-1}\mathbf{A}']}{|\boldsymbol{N} + \boldsymbol{S}|^{1/2}}. \tag{56}$$

The coupling matrices ${}_{\pm 2}\boldsymbol{W}$ can be computed quickly for an azimuthally symmetric patch of sky, as described in Appendix C, and modes with different m decouple. The problem is therefore tractable. However it is not nearly so simple to find a maximum likelihood estimate of the magnetic amplitude, and in general there will be complicated correlations between the recovered power spectra. By using the lossy projection in the previous sections we have essentially shown that this likelihood function is 'nearly' separable. Making it separable costs something in terms of lost information, but it significantly simplifies the problem. Using the projected variables also reduces the size of the matrices, so performing the matrix inversions is significantly faster.

If the B signal is determined to be negligible one would want to apply an efficient nearly-lossless method to estimate the electric power spectrum C_l^{EE} (or to do parameter estimation), so we now consider the case when one polarization type is absent. If the B signal can be neglected we have

$$\begin{pmatrix} \tilde{\mathbf{E}} \\ i\tilde{\mathbf{B}} \end{pmatrix} \approx \begin{pmatrix} \boldsymbol{W}_+ \\ \boldsymbol{W}_- \end{pmatrix} \mathbf{E} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^\dagger\mathbf{E}, \tag{57}$$

where we have done a singular value decomposition as before so that we can find the well determined linear combinations of the E_{lm}:

$$\mathbf{x}_E \equiv \tilde{\boldsymbol{D}}^{-1/2}\tilde{\boldsymbol{U}} \begin{pmatrix} \tilde{\mathbf{E}} \\ i\tilde{\mathbf{B}} \end{pmatrix} \approx \tilde{\boldsymbol{D}}^{1/2}\tilde{\boldsymbol{V}}^\dagger\mathbf{E}. \tag{58}$$

The matrices one has to invert to do a likelihood analysis are now one half the size of those in the optimal method when both polarization types are present, and so the problem is numerically much faster. However isotropic noise no longer gives the simple diagonal noise covariance, though this can always be rectified by using $\boldsymbol{N}^{-1/2}\mathbf{x}_E$. In practice a nearly optimal method would probably be more appropriate using only the unprojected variables $\tilde{\boldsymbol{D}}_+^{-1/2}\tilde{\boldsymbol{U}}_+^\dagger\tilde{\mathbf{E}}$, where $\tilde{\boldsymbol{D}}_+$ and $\tilde{\boldsymbol{U}}_+$ were defined in Sec. III. These variables have diagonal noise properties like the $\mathbf{E}_W$ for isotropic noise, and the computational saving may be significant when analysing high resolution polarization maps. We have checked numerically that including $\tilde{\mathbf{B}}$ in the analysis gains very little even for low $l_{\max}$ and small patches. For large area surveys at high resolution the information loss will probably be negligible.

There are exactly equivalent relations for the well determined magnetic variables in the case when E vanishes. This case is of little practical interest, since the E signal would have to be removed to within the magnitude of the B signal, and this is impossible on an incomplete sky since the two are not unambiguously distinguishable without accurate

boundary data. However, supposing that E could be removed is useful theoretically as we can then compute the best obtainable magnetic signal to compare with what we obtain using our projected variables. The information lost due to the projection depends on the cosmology. Models with reionization have more power on large scales and a greater fraction of the power is lost due to removal of the boundary terms. For our toy model of the Planck satellite we find that the amplitude that could be detected at given significance and probability is reduced by about 30 per cent by the projection for a cosmology with reionization at $z = 6.5$, but only by 2 per cent for a zero reionization model. In the reionization model one is loosing a lot of the additional information in the low multipoles that in the absence of the projection would have high signal to noise. The net result is that the reionization model has an only slightly higher chance of giving a tensor detection despite having more large scale power. By using the unprojected variables and incorporating the expected electric polarization contamination as an extra noise term one can approximately halve this loss. The lossless result is compared to the realistic projected result for general circular sky patches in Fig. 7 for an observation with much higher sensitivity. The cost we incur by using the non-optimal method in terms of slightly larger error bars on the B signal, or a less powerful test of detection at a given significance, is small for large survey areas though it does increase for small sky patches. For these sensitive observations the electric signal is much larger than the magnetic signal and essentially nothing is lost by performing the projection rather than including electric contamination as a large extra noise.

To make a detection of the magnetic signal on such a small sky patch with the planned long duration balloon observations the tensor/scalar ratio would need to be significantly larger than one, which is too large to be allowed by the current temperature anisotropy observations [28]. Of course, seeing if there is only a small magnetic signal is an important consistency check for current models with low tensor/scalar ratio to pass.

One simple way to reduce the information loss in our method would be to use data objects that include not only the surface integrals $I_{\pm 2,W}$, but also those parts of the boundary terms in Eqs. (7) and (8) that do not depend on $\eth P$ on the boundary. Such objects would separate electric and magnetic polarization exactly if the scalar window functions were constructed to vanish on the boundary. The problem of producing a non-redundant set of such windows could be tackled with a simple variant of the harmonic-based method presented in Sec. III. The additional boundary contribution would cancel that part of $\boldsymbol{W}_-$ that couples to the normal derivative of the window function on the boundary, leaving a single non-zero singular value (for $|m| > 0$) to project out. The net effect would be that for azimuthal patches we would gain one extra variable per m for $|m| > 1$, though the noise properties of these variables would not be as simple as if the line integrals were not included, and the problem of performing line integrals with pixelized data is non-trivial. For reionization models (which have significant large scale power) the reduction in information loss may be worth the effort required to overcome these obstacles, though a full analysis with the non-separated variables would probably work better.

VI. CONCLUSION

We have considered the problem of producing statistically independent measures of the electric and magnetic polarization from observations covering only a portion of the sky. Although the separation of the polarization field into electric and magnetic modes is not unique in the presence of boundaries, we have shown how to construct window functions that are guaranteed to probe separately the electric and magnetic polarization exactly over azimuthally-symmetric patches of the sky. We presented a harmonic-based method for efficient construction of the windows that automatically removes redundancy due to the finite sky coverage. In addition, our window functions return separated electric and magnetic variables that have very simple diagonal noise correlations for idealized noise on the polarization map. For azimuthal patches separating the electric and magnetic polarization comes at the cost of losing two pieces of information per m mode, or roughly twice the number of pixels of area $(\pi/l_{\rm max})^2$ on the boundary of the patch. For large patches this information loss is small unless there is large scale power due to reionization, but for smaller patches it can be more severe due to the limited support of the high m spin-weight 2 harmonics in the patch. Although we have proved that our method gives exact separation for azimuthal patches, the harmonic-based construction should produce window functions that give approximate separation for arbitrarily shaped patches with similar information loss to the azimuthal case.

We showed how the variables constructed from our window functions could be used to constrain the amplitude of the magnetic signal without contamination from the much larger electric signal. For the first time, we made predictions for the tensor amplitude that Planck should be able to detect taking proper account of excluding the galactic region. If other non-negligible foregrounds can be removed using the other frequency channels, Planck should be able to detect the magnetic signal predicted by some simple inflationary models. For less sensitive observations, our window functions should nevertheless be useful to set upper limits on the magnetic signal, and may also aid the identification of systematic effects in the instrument or analysis pipeline.

If the magnetic signal is shown to be consistent with noise then we showed how one can use all the well determined

polarization pseudo-multipoles to analyse the electric polarization power spectrum without loss of information. The analysis using these variables is no more complicated than the analysis of temperature anisotropies using cut-sky orthogonalized scalar harmonic functions [10, 11].

We have only considered isotropic noise here, however, as long as the noise is azimuthally symmetric the separation of m modes will still work, and the problem remains computationally tractable though rather less simple. In practice, there will be several other complications in real-life CMB polarimetry observations that will impact on the map-making and subsequent analysis stages. Further careful investigation of the propagation of instrument effects such as beam asymmetries, straylight, cross-polar contamination, and pointing instabilities through the map-making stage will be required before the programme for analysing magnetic polarization outlined in this paper will be realizable.

Acknowledgments

We acknowledge use of the LAPACK package[2] for performing the matrix decompositions. We acknowledge the support of PPARC via a PPTC Special Program Grant for this work. AC acknowledges a PPARC Postdoctoral Fellowship. We also thank PPARC and HEFCE for support of the COSMOS facility.

APPENDIX A: SPIN RAISING AND LOWERING OPERATORS

In general a spin-weight s quantity ${}_s\eta$ is defined over a two-dimensional Riemannian manifold with respect to an orthonormal diad field $\{\mathbf{e}_1, \mathbf{e}_2\}$. The local freedom in the choice of diad amounts to the transformations

$$\mathbf{e}_\pm \equiv \mathbf{e}_1 \pm i\mathbf{e}_2 \rightarrow e^{i\gamma}\mathbf{e}_\pm \tag{A1}$$

of the (complex) null vectors $\mathbf{e}_\pm$. A quantity is defined to be of spin-weight s if under the transformation (A1) ${}_s\eta \rightarrow {}_s\eta e^{is\gamma}$. To every spin-weight s object ${}_s\eta$ we can associate a (complex) symmetric trace-free, rank-$|s|$ tensor $\eta_{A_s} \equiv \eta_{a_1\ldots a_s}$: for $s \geq 0$,

$$\eta^{A_s} \equiv 2^{-s}{}_s\eta e_-^{A_s}, \tag{A2}$$

where the irreducible tensor product $e_-^{A_s} \equiv e_-^{a_1}\ldots e_-^{a_s}$. The inverse relation is

$${}_s\eta = \eta_{A_s}e_+^{A_s}. \tag{A3}$$

For $s < 0$ we define $\eta^{A_{|s|}} \equiv 2^{-|s|}{}_s\eta e_+^{A_{|s|}}$. The spin raising and lowering operators $\eth$ and $\bar{\eth}$ are defined by the null diad components of the covariant derivatives of $\eta_{A_{|s|}}$:

$$\eth\,{}_{\pm|s|}\eta = -(e_+^c\nabla_c\eta_{A_{|s|}})e_\pm^{A_{|s|}}, \tag{A4}$$

$$\bar{\eth}\,{}_{\pm|s|}\eta = -(e_-^c\nabla_c\eta_{A_{|s|}})e_\pm^{A_{|s|}}. \tag{A5}$$

(The minus signs are conventional).

In CMB polarimetry we are concerned with fields defined over the sphere, in which case the transformation in Eq. (A1) corresponds to a *left*-handed rotation of the diad about the outward normal $\hat{\mathbf{r}}$. Choosing the orthonormal diad to be aligned with the coordinate basis vectors $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ of a spherical polar coordinate system, we have $e_\pm^a\nabla_a e_\pm^b = \cot\theta e_\pm^b$ and $e_\mp^a\nabla_a e_\pm^b = -\cot\theta e_\pm^b$. It follows that for this choice of diad the spin raising and lowering operators reduce to

$$\eth\,{}_s\eta = -\sin^s\theta(\partial_\theta + i\csc\theta\partial_\phi)(\sin^{-s}\theta\,{}_s\eta), \tag{A6}$$

$$\bar{\eth}\,{}_s\eta = -\sin^{-s}\theta(\partial_\theta - i\csc\theta\partial_\phi)(\sin^s\theta\,{}_s\eta). \tag{A7}$$

An elegant interpretation of the spin raising and lowering operators on the sphere can be obtained by considering spin-weight s objects ${}_s\eta(\theta,\phi,\psi)$ defined on a diad at (position-dependent) angle ψ to the coordinate directions, so that

$${}_s\eta(\theta,\phi,\psi) = {}_s\eta(\theta,\phi)e^{is\psi}, \tag{A8}$$

[2] http://www.netlib.org/lapack/

where ${}_s\eta(\theta,\phi)$ is defined on $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$. In this case, the spin raising and lowering operators can be related to the angular momentum operators for a rigid body [16]. Working in a representation where the orientation of the body is specified in terms of Euler angles $(\phi,\theta,-\psi)$[3], the angular momentum operators on the *body-fixed axes* take the form

$$K_z = i\partial_\psi, \tag{A9}$$

$$K_\pm = e^{\pm i\psi}(\pm\partial_\theta + i\csc\theta\partial_\phi + i\cot\theta\partial_\psi), \tag{A10}$$

where $K_\pm \equiv K_x \pm iK_y$. These operators satisfy the commutation relations

$$[K_z, K_\pm] = \mp K_\pm, \qquad [K_+, K_-] = -2K_z, \tag{A11}$$

so that $K_\pm$ are lowering/raising operators with respect to the eigenvalues of K_z. Note that the signs in these commutation relations are different from those for angular momentum operators on a fixed frame since on the body-fixed axes we have $[K_x, K_y] = -iK_z$ [30]. The action of the spin raising and lowering operators can then be formulated in terms of the angular momentum operators as

$$K_+{}_s\eta(\theta,\phi,\psi) = -e^{i(s+1)\psi}\,\eth\,{}_s\eta(\theta,\phi), \tag{A12}$$

$$K_-{}_s\eta(\theta,\phi,\psi) = +e^{i(s-1)\psi}\,\overline{\eth}\,{}_s\eta(\theta,\phi). \tag{A13}$$

Several useful results for the spherical raising and lowering operators follow from the commutation relations (A11). [Similar relations on a general manifold can be found from Eqs. (A4) and (A5).] For a spin-weight s quantity defined on the coordinate basis,

$$(\,\overline{\eth}\,\eth - \eth\,\overline{\eth}\,)_s\eta(\theta,\phi) = 2s\,{}_s\eta(\theta,\phi), \tag{A14}$$

$$[\,\eth\,\overline{\eth} - s(s-1)]_s\eta(\theta,\phi) = (\nabla^2 - s^2\csc^2\theta + 2is\cot\theta\,\csc\theta\,\partial_\phi)_s\eta(\theta,\phi), \tag{A15}$$

where we have used $K_+K_- = K^2 - K_z - K_z^2$ to derive the last identity. Applying these relations repeatedly to a spin-weight 0 quantity we find

$$\eth\,\eth\,\overline{\eth}\,\overline{\eth}\,{}_0\eta(\theta,\phi) = \overline{\eth}\,\overline{\eth}\,\eth\,\eth\,{}_0\eta(\theta,\phi) = (\nabla^2+2)\nabla^2\,{}_0\eta(\theta,\phi), \tag{A16}$$

and the useful relation

$$(\,\overline{\eth}\,\eth^s - \eth^s\,\overline{\eth}\,)_0\eta(\theta,\phi) = s(s-1)\,\eth^{s-1}{}_0\eta(\theta,\phi). \tag{A17}$$

Integral theorems

The integral of the spin-weight 0 quantity $\eth\,{}_{-1}\eta$ over some portion S of the two-dimensional manifold is determined by the integral around the boundary ∂S:

$$\int_S \mathrm{d}S\;\eth\,{}_{-1}\eta = \oint_{\partial S} {}_1\mathrm{d}l\;{}_{-1}\eta, \tag{A18}$$

where ${}_1\mathrm{d}l$ is the spin-one element of length around the boundary:

$${}_1\mathrm{d}l \equiv i\mathrm{d}l_a e^a_+. \tag{A19}$$

On the sphere in the spherical polar coordinate frame ${}_1\mathrm{d}l = i\mathrm{d}\theta - \sin\theta\,\mathrm{d}\phi$. Equation (A18) is the complex representation of Stokes' Theorem and the Divergence Theorem. An equivalent result holds for spin-weight one quantities by forming the complex conjugate of the above, with ${}_1\overline{\mathrm{d}l} = -i\mathrm{d}l_a e^a_-$.

The spin raising and lowering operators obey Leibnitz' rule, so there is a 'Green's identity'

$$P\,\eth\,\overline{\eth}\,Q - Q\,\overline{\eth}\,\eth\,P = \eth\,(P\,\overline{\eth}\,Q) - \overline{\eth}\,(Q\,\eth\,P), \tag{A20}$$

---

[3] Our convention for Euler angles (α,β,γ) follows Ref. [29], i.e., successive right-handed rotations by γ, β, and α about the z, y, and z-axes respectively. The use of ψ, which is minus the third Euler angle, as a configuration variable for the rigid body is necessary to relate the angular momentum operators directly to the spin raising and lowering operators with the (consistent) conventions we have adopted here.

where P and Q have definite spin weight. For PQ with spin-weight zero integrating over a surface gives the integral theorem

$$\int_S \mathrm{d}S(P\,\eth\,\overline{\eth}\,Q - Q\,\overline{\eth}\,\eth\,P) = \oint_{\partial S} {}_1\mathrm{d}l\; P\,\overline{\eth}\,Q - \oint_{\partial S} {}_1\overline{\mathrm{d}l}\; Q\,\eth\,P. \tag{A21}$$

A similar result is obtained using

$$P\,\eth\,\eth Q - Q\,\eth\,\eth P = \eth\,(P\,\eth Q - Q\,\eth P) \tag{A22}$$

for PQ of spin weight -2 which gives

$$\int_S \mathrm{d}S(P\,\eth\,\eth Q - Q\,\eth\,\eth P) = \oint_{\partial S} {}_1\mathrm{d}l\;(P\,\eth Q - Q\,\eth P), \tag{A23}$$

with a similar result for a spin-weight 2 quantity.

APPENDIX B: SPIN WEIGHTED SPHERICAL HARMONICS

The spin-weight s spherical harmonics ${}_sY_{lm}$ are defined in terms of the usual spherical harmonics Y_{lm} by

$${}_sY_{lm} \equiv \sqrt{\frac{(l-|s|)!}{(l+|s|)!}}\;\eth^s\,Y_{lm}, \tag{B1}$$

where $\eth^{\,-|s|} \equiv (-1)^s\,\overline{\eth}^{\,|s|}$. They are non-zero for $|s| \le l, |m| \le l$. By making use of Eqs. (A12) and (A13), and the properties of the $K_\pm$ operators when acting on Wigner D-matrices[4] [29, 31], it is straightforward to show that [21]

$$D^l_{-ms}(\phi,\theta,-\psi) = (-1)^m\sqrt{\frac{4\pi}{2l+1}}\,{}_sY_{lm}(\theta,\phi)e^{is\psi}. \tag{B2}$$

With the conventions adopted here, ${}_sY^*_{lm} = (-1)^{s+m}\,{}_{-s}Y_{l(-m)}$. Orthonormality of the spin weight harmonics over the full sphere,

$$\int_{4\pi} \mathrm{d}S\;{}_sY_{lm}\,{}_sY^*_{l'm'} = \delta_{ll'}\delta_{mm'}, \tag{B3}$$

follows from the orthogonality of the D-matrices over the SO(3) group manifold.

Applying K_+K_- to Eq. (B2), and using Eqs. (A12) and (A13), one can show that the spin weight harmonics satisfy the differential equation

$$\eth\;\overline{\eth}\;{}_sY_{lm} = [s(s-1) - l(l+1)]\;{}_sY_{lm}. \tag{B4}$$

[An alternative proof of this result follows from Eq. (A17).] The spin weighted harmonics are separable and can be written as

$${}_sY_{lm}(\theta,\phi) = {}_s\lambda_{lm}(\cos\theta)e^{im\phi}. \tag{B5}$$

The ${}_s\lambda_{lm}$ satisfy the self-adjoint equation

$$[(1-x^2)\,{}_s\lambda'_{lm}]' - \frac{m^2+s^2+2msx}{1-x^2}\,{}_s\lambda_{lm} = -l(l+1)\,{}_s\lambda_{lm}, \tag{B6}$$

where a prime denotes differentiation with respect to $x = \cos\theta$.

---

[4] Our conventions for the D-matrices follow Refs. [29, 31]. We adopt the Condon-Shortley phase for the spherical harmonics, which differs from that used by Goldberg et al. [16] by a factor of $(-1)^m$.

The ${}_s\lambda_{lm}$ can be evaluated recursively for $m \geq |s|$ starting with

$${}_s\lambda_{mm}(x) = (-2)^{-m}\sqrt{\frac{(2m+1)!}{4\pi(m+s)!(m-s)!}}(1-x)^{(m+s)/2}(1+x)^{(m-s)/2}, \tag{B7}$$

and the recursion relation (derived from standard results for the Wigner D-matrices; see e.g. Ref. [31])

$${}_s\lambda_{lm} = \left(x + \frac{sm}{l(l-1)}\right) C_{slm}\, {}_s\lambda_{(l-1)m} - \frac{C_{slm}}{C_{s(l-1)m}}\, {}_s\lambda_{(l-2)m}, \tag{B8}$$

where

$$C_{slm} \equiv \sqrt{\frac{l^2(4l^2-1)}{(l^2-m^2)(l^2-s^2)}}. \tag{B9}$$

The harmonics for $m \leq -|s|$ can be obtained from ${}_s\lambda_{l(-m)} = (-1)^{s+m}\, {}_{-s}\lambda_{lm}$. A straightforward way to evaluate ${}_s\lambda_{lm}$ for $|m| < |s|$ is to compute ${}_{\pm n}\lambda_{l|s|}$ for $0 \leq n < |s|$ and then use the symmetry ${}_s\lambda_{lm} = (-1)^{m+s}\, {}_m\lambda_{ls}$. Further useful results for the spin weight harmonics can be found in Refs. [4, 22], and expressions for the spin-weight ± 2 harmonics in terms of the associated Legendre functions are given in Ref. [3].

APPENDIX C: OVERLAP INTEGRALS

Setting $P = {}_sY^*_{l'm'}$ and $Q = {}_sY_{lm}$ in Eq. (A21), and using the differential equation (B4) to simplify the integrand on the left-hand side, we have

$$[l'(l'+1) - l(l+1)] \int_S \mathrm{d}S\, {}_sY^*_{l'm'}\, {}_sY_{lm} = \oint_{\partial S} {}_1\mathrm{d}l\, {}_sY^*_{l'm'}\, \overline{\eth}\, {}_sY_{lm} - \oint_{\partial S} {}_1\overline{\mathrm{d}l}\, {}_sY_{lm}\, \eth\, {}_sY^*_{l'm'}. \tag{C1}$$

This expresses the overlap integral for $l \neq l'$ in terms of a line integral around the boundary of S.

For an azimuthally symmetric surface the integrals are particularly straightforward as the spin-weight harmonics with different m are orthogonal over the patch. The overlap integral for the same m but different l can be determined from Eq. (C1) to be

$$[l(l+1) - l'(l'+1)] \int_a^b \mathrm{d}x\, {}_s\lambda_{l'm}\, {}_s\lambda_{lm} = (1-x^2)[{}_s\lambda'_{l'm}\, {}_s\lambda_{lm} - {}_s\lambda'_{lm}\, {}_s\lambda_{l'm}]_a^b. \tag{C2}$$

[This result also follows directly from Eq. (B6).] The derivatives can be removed while maintaining homogeneity in the spin weight by using

$$l(1-x^2)\, {}_s\lambda'_{lm} = -(sm + l^2 x)\, {}_s\lambda_{lm} + \frac{(2l+1)l}{C_{slm}}\, {}_s\lambda_{(l-1)m} \tag{C3}$$

to write the $l' \neq l$ integral as

$$A^{sm}_{ll'} \equiv 2\pi \int_a^b \mathrm{d}x\, {}_s\lambda_{l'm}\, {}_s\lambda_{lm} = \frac{2\pi}{(l+l'+1)(l-l')} \left[\left(x - \frac{sm}{ll'}\right)(l-l')\, {}_s\lambda_{l'm}\, {}_s\lambda_{lm} + \frac{2l'+1}{C_{sl'm}}\, {}_s\lambda_{(l'-1)m}\, {}_s\lambda_{lm} - \frac{2l+1}{C_{slm}}\, {}_s\lambda_{(l-1)m}\, {}_s\lambda_{l'm} \right]_a^b. \tag{C4}$$

Note that $A^{sm}_{ll'} = A^{(-s)(-m)}_{ll'}$. The $l = l' \neq |m|$ integrals can be evaluated recursively using

$$A^{sm}_{ll} = A^{sm}_{(l-1)(l-1)} + \frac{C_{slm}}{C_{s(l+1)m}} A^{sm}_{(l+1)(l-1)} - \frac{C_{slm}}{C_{s(l-1)m}} A^{sm}_{l(l-2)} + \frac{2sm}{l(l^2-1)} C_{slm} A^{sm}_{l(l-1)}, \tag{C5}$$

which follows from Eq. (B8). For $|m| > |s|$, the starting values $A^{sm}_{|m||m|}$ can be obtained from

$$A^{sm}_{mm} = A^{s(m-1)}_{(m-1)(m-1)} + \frac{2\pi[x_s \lambda^2_{mm}]_a^b}{2m+1} + \frac{s}{\sqrt{(2m+1)(m^2-s^2)}} A^{s(m-1)}_{m(m-1)}, \tag{C6}$$

$$A^{s(-m)}_{mm} = A^{s(-m+1)}_{(m-1)(m-1)} + \frac{2\pi[x_s \lambda^2_{m(-m)}]_a^b}{2m+1} - \frac{s}{\sqrt{(2m+1)(m^2-s^2)}} A^{s(-m+1)}_{m(m-1)} \qquad (m > |s|), \tag{C7}$$

which follow from the explicit form of ${}_s\lambda_{mm'}$ given in Eq. (B7). For $l' = l = m = |s|$ the integrals have the analytic solutions

$$A^{mm}_{mm} = \left[\left(\frac{x-1}{2}\right)^{2m+1}\right]^b_a, \qquad A^{(-m)m}_{mm} = \left[\left(\frac{x+1}{2}\right)^{2m+1}\right]^b_a. \tag{C8}$$

For $|m| \leq |s|$, the starting values $A^{sm}_{|s||s|}$ for the recursion in Eq. (C5) can be obtained from the symmetry $A^{sm}_{ll'} = A^{ms}_{ll'}$ and the recursion relations (C6) and (C7).

If one is also generating the spin-weight zero overlap integrals for analysing the temperature field an alternative approach is to use Eq. (A23) to relate the integrals of the spin-weight two and zero harmonics. In general integrals for different spin weights can be related by

$$\begin{aligned} k_{l'(s-1)}k_{l's}\int_S \mathrm{d}S\, {}_sY^*_{lm}\, {}_sY_{l'm'} \;-\; & k_{l(s-1)}k_{ls}\int_S \mathrm{d}S\, {}_{s-2}Y^*_{lm}\, {}_{s-2}Y_{l'm'} \\ = & \oint_{\partial S} {}_1\mathrm{d}l\, \left(k_{l'(s-1)}\, {}_sY^*_{lm}\, {}_{s-1}Y_{l'm'} + k_{ls}\, {}_{s-2}Y_{l'm'}\, {}_{s-1}Y^*_{lm}\right), \end{aligned} \tag{C9}$$

where $k_{ls} \equiv \sqrt{l(l+1) - s(s-1)}$, and we have used the results

$$\eth\, {}_sY_{lm} = k_{l(s+1)}\, {}_{s+1}Y_{lm} \tag{C10}$$

$$\overline{\eth}\, {}_sY_{lm} = -k_{ls}\, {}_{s-1}Y_{lm} \tag{C11}$$

which follow from Eq. (B1). For $s = 0$ and $s = 2$ one can obtain the spin ± 2 integrals in terms of the spin zero integrals. The spin zero integrals are computed using the above relations with $s = 0$, in agreement with the relations given in Ref. [32].

For a small patch of sky a large number of the overlap integrals are going to be very close to zero. This makes sense intuitively, and is easy to see more quantitatively. From the differential equation (B6) the character of the harmonics changes from oscillatory to decaying at the point where

$$\frac{m^2 + s^2 + 2msx}{1 - x^2} = l(l+1). \tag{C12}$$

For x nearer the poles than the critical value the harmonics become very small. For a sky patch extending from the north pole to x, for large l we see that for m^2 larger than $m^2 \approx l^2(1-x^2)$ the harmonics will be small within the patch. Hence the overlap integrals with $m^2 > \min(l^2, l'^2)(1-x^2)$ will be small, corresponding to one (or both) of the harmonics being localized out of the sky patch.

APPENDIX D: STATISTICS OF WEAK SIGNAL DETECTION

After obtaining a vector of observed data $\mathbf{B}_W$ two questions one might ask are: (1) What is the probability that the signal is just noise?; (2) Given the signal is not noise, what can we say about the amplitude r? The likelihood function (48) encapsulates all the information in the data concerning the amplitude r, and gives the posterior probability distribution of r after multiplying by the prior. Question (1) is really asking for a comparison of two models, one in which $r = 0$ (or $r < \epsilon$ where ϵ is small), and one in which $r > 0$. Given there is a prior probability p that the signal is pure noise ($r = 0$) and probability $1 - p$ that $r > 0$, distributed with the normalized prior probability distribution $f(r)$, the posterior probability that $r = 0$ is given by

$$P(r = 0|\mathbf{B}_W) = \frac{pL(\mathbf{B}_W|r=0)}{pL(\mathbf{B}_W|r=0) + (1-p)\int_{0_+}^{\infty} \mathrm{d}r L(\mathbf{B}_W|r) f(r)}. \tag{D1}$$

The posterior probability tells us the probability that the signal is pure noise once we have a particular set of data. However, it doesn't immediately tell us that we might expect to obtain from a given observation, just what we know once the observation has been performed. The posterior distribution can depend strongly on the prior.

The posterior probability is a monotonic function of the Bayes factor

$$t \equiv \frac{\int_{0_+}^{\infty} \mathrm{d}r L(\mathbf{B}_W|r) f(r)}{L(\mathbf{B}_W|r=0)}. \tag{D2}$$

In classical hypothesis testing the likelihood ratio $L(\mathbf{B}_W|r')/L(\mathbf{B}_W|r=0)$ is the most powerful test statistic for distinguishing a model with $r = r'$ from one with $r = 0$ (for details see Ref. [33]). The largest fraction $1-\alpha$ of the values of the likelihood ratio under the null-hypothesis ($r = 0$) determine a region, which, if the observed ratio falls in it, rejects the null hypothesis at significance α. The distribution of the likelihood ratio is straightforward to compute using Monte-Carlo techniques. In general we do not have some fixed alternative hypothesis $r = r'$; one possibility is to use the the statistic t, formed by marginalizing over some prior, in place of the likelihood ratio. Often the prior is fixed to $\delta(r-\hat{r})$, where $\hat{r}$ is the maximum likelihood estimate of r, so the test reduces to a likelihood ratio test with the alternative hypothesis $r = \hat{r}$.

By the time magnetic polarization comes to be observed we should have some prior information about the tensor amplitude from accurate measurement of the temperature power spectrum. However at this point we do not have very useful information about the prior distribution, and to consider the possible results of future experiments it is useful to assume no information. The maximum entropy prior (the uniform prior) is probably not appropriate in this case — do we really think that all values of r (below a certain bound) are equally likely rather than, say, all values of $\ln r$ being equally likely? The uniform prior gives radically different answers depending on the choice of variable. The Jeffreys prior

$$f(r) \propto \left\langle -\frac{\partial^2}{\partial r^2} \ln L(r|\mathbf{B}_W) \right\rangle^{1/2} = \tfrac{1}{2}\mathrm{tr}\left[(\boldsymbol{N}+r\boldsymbol{S})^{-1}\boldsymbol{S}(\boldsymbol{N}+r\boldsymbol{S})^{-1}\boldsymbol{S}\right]^{1/2}, \tag{D3}$$

is reparameterization invariant [34], and therefore does not suffer from this problem. The Jeffreys prior goes like $1/(1+r)$ if $\boldsymbol{N} = \boldsymbol{S}$ and in general is improper (does not have a finite integral), though this is not a problem for evaluating test statistics since the integral of the product with the likelihood function will be finite. It is also not a true prior in the sense that it depends on what data is going to be collected, but this is really a good thing as it concentrates the prior probability in the region where we need it in order to obtain a detection. Since the prior is not localized the posterior probability of the null-hypothesis will depend on where one cuts off the prior. However if one only ever compares the Bayes factors t the cut-off is not very important since the likelihood function will be localized. For this reason performing classical hypothesis tests using t is rather more independent of the prior information than considering the values of the posterior probability, though clearly it is the latter which is rigorous and contains all the available information.

Classical hypothesis tests are useful for assessing probability of getting a detection at a given significance, though it should be remembered that getting a detection at 99 per cent confidence does *not* mean a the probability of 1 per cent that the signal is pure noise. However for sensible priors there will be a close correspondence; classical and Bayesian techniques agree that high values of t correspond to the null-hypothesis being less likely. Using classical hypothesis tests to compute the probability of getting a detection at a given significance for a given true r is significantly simpler than computing and interpreting the corresponding distribution of the posterior probability distributions.

For full sky surveys there are a large number of statistically-independent magnetic variables, which can be thought of as the eigenvectors of $r\boldsymbol{N}^{-1/2}\boldsymbol{S}\boldsymbol{N}^{-1/2}$. Those variables corresponding to large eigenvalues have high expected signal to noise and are most useful for obtaining detections. If $\mathbf{B}_W$ has dimension n there will be n independent variables. If the eigenvalues of $r\boldsymbol{N}^{-1/2}\boldsymbol{S}\boldsymbol{N}^{-1/2}$ are distributed fairly uniformly, and we consider expected chi-squared detections of order one sigma, all the eigenvalues will be $\sim 1/\sqrt{n}$ and we have $r\boldsymbol{S} \ll \boldsymbol{N}$. Using a second order approximation we then have

$$(\boldsymbol{N}+r\boldsymbol{S})^{-1} \approx (\boldsymbol{I} - r\boldsymbol{N}^{-1}\boldsymbol{S} + r^2\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S})\boldsymbol{N}^{-1}, \tag{D4}$$

and

$$\ln|\boldsymbol{N}+r\boldsymbol{S}| = \ln|\boldsymbol{N}| + \ln|\boldsymbol{I}+r\boldsymbol{N}^{-1}\boldsymbol{S}| \approx \ln|\boldsymbol{N}| + \mathrm{tr}(r\boldsymbol{N}^{-1}\boldsymbol{S} - \tfrac{1}{2}r^2\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S}). \tag{D5}$$

In this approximation the entire likelihood distribution is simple to compute since the matrix manipulations only need be performed once, and is of the Gaussian form

$$L(\mathbf{B}_W|r) = L(\mathbf{B}_W|r=0)e^{\hat{r}^2/2\sigma^2}e^{-(r-\hat{r})^2/2\sigma^2}, \tag{D6}$$

where

$$1/\sigma^2 = \mathbf{B}_W^\dagger\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\mathbf{B}_W - \tfrac{1}{2}\mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S}), \tag{D7}$$

and the maximum likelihood estimate for r is

$$\hat{r} \approx \tfrac{1}{2}\sigma^2\left[\mathbf{B}_W^\dagger\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\mathbf{B}_W - \mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S})\right]. \tag{D8}$$

The maximum likelihood estimate for r is only weakly biased for small signal to noise.

Consider fixing the signal hypothesis to $r = \hat{r}$, which gives the largest likelihood ratio possible for any prior. The likelihood ratio is then a monotonic function of the test statistic

$$\nu' = \hat{r}/\sigma \equiv \frac{\mathbf{B}_W^\dagger \boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\mathbf{B}_W - \mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S})}{\sqrt{4\mathbf{B}_W^\dagger \boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\mathbf{B}_W - 2\mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S})}}. \tag{D9}$$

The quantity $\hat{r}/\sigma$ gives the number of standard deviations the maximum likelihood is from pure noise, which is an good intuitive measure of the number of 'sigmas' at which the magnetic signal has been detected. If any eigenvalues become large enough then the small r approximation will fail and one needs to compute the full likelihood ratio to obtain optimal results. The ν' statistic is a function only of the vector $\boldsymbol{N}^{-1/2}\mathbf{B}_W$ and the matrix $\boldsymbol{N}^{-1/2}\boldsymbol{S}\boldsymbol{N}^{-1/2}$, so working in the frame in which the matrix is diagonal the statistic is simple to compute. Note that σ^2 and $\hat{r}$ can be negative and correspond to a non-detection.

The second order approximation can also be used to speed up computing the full likelihood function for integration against a prior. In the diagonal frame it is straightforward to identify any modes with high signal to noise — if there are some then the likelihood for these modes can be computed exactly. The likelihood for the remaining low signal to noise modes can be computed quickly using the approximation in which the likelihood is Gaussian. Multiplying these together allows the full likelihood function to be computed, and hence the the probability distribution of t. Whilst this is slower than using the ν' statistic it is a useful check, and may be essential when there are very high signal to noise modes.

For good detections the prior is not very important as long as it is not small over the bulk of the likelihood. If the prior is approximately constant over most of the likelihood integral, the second order approximation is valid, and $\hat{r}/\sigma \gg 0$, then $t \sim e^{\hat{r}^2/2\sigma^2} g(\hat{r}, \sigma)$ where $g(\hat{r}, \sigma)$ depends weakly on $\hat{r}$ and σ compared to the exponential. In this case t depends almost entirely (and monotonically) on the value of $\nu' = \hat{r}/\sigma$, which is why ν' is a good statistic to use. We found that detection probabilities computed using full likelihood results with the Jeffreys prior tend to agree very closely with those computed using ν'.

The advantage of using ν' rather than computing the full likelihood distribution is that it is computationally significantly simpler and faster, which is useful, though by no means essential, for performing accurate Monte-Carlo computations. There is no problem computing using the full likelihood distribution exactly from single samples of actual observed data.

Busting the null-buster

In the limit that $r \ll 1$ the likelihood function can be approximated by the first two terms in its Taylor series in r. The likelihood ratio for any given (small) r is then a monotonic function of $\partial_r \ln L(\mathbf{B}_W|r)|_{r=0}$, which is independent of r and therefore provides the uniformly most powerful test. For Gaussian signal and noise it is proportional to $2\hat{r}/\sigma^2$. After dividing by the root of the variance in the null hypothesis $\langle (2\hat{r}/\sigma^2)^2 \rangle|_{r=0} = 2\mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S})$ we obtain the quantity

$$\nu \equiv \frac{\mathbf{B}_W^\dagger \boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\mathbf{B}_W - \mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S})}{\sqrt{2\mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S})}}, \tag{D10}$$

which is the optimal[5] quadratic 'null-buster' statistic introduced to the CMB literature in Ref. [35]. In the limit in which all the eigenvalues of $r\boldsymbol{N}^{-1/2}\boldsymbol{S}\boldsymbol{N}^{-1/2}$ tend to zero this is the optimal test statistic. In general it is not — a signal that can be detected will violate this assumption. For a very large number of approximately equal eigenvalues the value of σ will approximate $\langle 1/\sigma^2 \rangle^{-1/2} = [\frac{1}{2}\mathrm{tr}(\boldsymbol{N}^{-1}\boldsymbol{S}\boldsymbol{N}^{-1}\boldsymbol{S})]^{-1/2}$, the Fisher curvature for small r. In this limit the null-buster remains a good statistic and $\nu'/\nu \to 1$. However in general ν' performs significantly better when the eigenvalues are distributed less evenly, or when there are not that many eigenvalues. As the signal to noise increases ν' also performs much less sub-optimally than the null-buster.

For the case of magnetic polarization observations the ν' statistic outperforms the null-buster for a wide range of patch sizes in realistic reionization models. The large scale magnetic signal coming from low redshift ($z < 10$)

[5] It is the *quadratic* statistic which gives the maximal *expected* detection in units of the *expected standard deviation* under the null hypothesis. None of these properties are required or even especially desired (we are more interested in getting detections at high *significance* with high *probability*).

reionization gives a small number of modes with relatively high signal to noise (see Fig. 4), and the conditions under which the null-buster is a good statistic are therefore not satisfied. The qualitative reason that the null-buster performs poorly is that the position of the maximum and the curvature of the likelihood function are correlated, so dividing the actual maximum by the expected curvature does not give you an accurate measure of the number of 'sigmas' from zero for a particular observation. This makes the null hypothesis distribution unnecessarily broad at large values, and therefore makes it harder to rule out the null hypothesis with good significance. The ν' statistic has a much sharper distribution than the null-buster (which has a distribution similar to chi-squared) in the alternative hypothesis, and the value of ν' corresponds much more closely to the significance (measured in Gaussian-like 'sigmas') of the result.

A few points can be made in the null-buster's defence. Firstly it is slightly easier to compute than ν'. Secondly, since we motivated the ν' by assuming a Gaussian signal it is conceivable that the null-buster could perform better with certain non-Gaussian signals. Lastly, the null-buster is quadratic which makes it easy to calculate the mean and variance analytically. However it is clear that with Gaussian signals using the null-buster is in general significantly sub-optimal.

[1] M. Bucher, K. Moodley, and N. Turok, Phys. Rev. Lett. in press (2001), astro-ph/0012141.
[2] W. Hu, U. Seljak, M. White, and M. Zaldarriaga, Phys. Rev. D **57**, 3290 (1998), astro-ph/9709066.
[3] M. Kamionkowski, A. Kosowsky, and A. Stebbins, Phys. Rev. D **55**, 7368 (1997), astro-ph/9611125.
[4] M. Zaldarriaga and U. Seljak, Phys. Rev. D **55**, 1830 (1997), astro-ph/9609170.
[5] T. Chiueh and C.-J. Ma (2001), astro-ph/0101205.
[6] M. Zaldarriaga (2001), astro-ph/0106174.
[7] W. Hu (2001), astro-ph/0108090.
[8] J. Guzik, U. Seljak, and M. Zaldarriaga, Phys. Rev. D **62**, 043517 (2000), astro-ph/9912505.
[9] K. Benabed, F. Bernardeau, and L. van Waerbeke, Phys. Rev. D **63**, 043501 (2001), astro-ph/0003038.
[10] K. M. Górski, Astrophys. J. Lett. **430**, 85 (1994), astro-ph/9403066.
[11] D. J. Mortlock, A. D. Challinor, and M. P. Hobson (2000), astro-ph/0008083.
[12] M. Tegmark and A. de Oliveira-Costa, Phys. Rev. D **64**, 063001 (2001), astro-ph/0012120.
[13] N. Kaiser, Astrophys. J. **388**, 272 (1992).
[14] A. Stebbins (1996), astro-ph/9609149.
[15] R. G. Crittenden, P. Natarajan, U. Pen, and T. Theuns (2000), astro-ph/0012336.
[16] J. N. Goldberg, A. J. Macfarlane, E. T. Newman, F. Rohrlich, and E. C. G. Sudarshan, J. Math. Phys. **8**, 2155 (1967).
[17] E. Newman and R. Penrose, J. Math. Phys. **7**, 863 (1966).
[18] G. H. Golub and C. F. van Loan, *Matrix operations* (The John Hopkins Univ. Press, Baltimore, 1988), 3rd ed.
[19] A. D. Challinor, D. J. Mortlock, F. van Leeuwen, A. N. Lasenby, M. P. Hobson, M. A. J. Ashdown, and G. P. Efstathiou (2001), in preparation.
[20] F. Couchot, J. Delabrouille, J. Kaplan, and B. Revenu, Astron. Astrophys. Suppl. **135**, 579 (1999), astro-ph/9807080.
[21] A. Challinor, P. Fosalba, D. Mortlock, M. Ashdown, B. Wandelt, and K. Górski, Phys. Rev. D **62**, 123002 (2000), astro-ph/0008228.
[22] K.-W. Ng and G.-C. Liu, Int. J. Mod. Phys. D **8**, 61 (1999), astro-ph/9710012.
[23] M. Zaldarriaga and U. Seljak, Phys. Rev. D **58**, 023003 (1998), astro-ph/9803150.
[24] J. Martin and D. J. Schwarz, Phys. Rev. D **62**, 103520 (2000), astro-ph/9911225.
[25] A. H. Jaffe, M. Kamionkowski, and L. Wang, Phys. Rev. D **61**, 083501 (2000), astro-ph/9909281.
[26] E. F. Bunn (2001), astro-ph/0108209.
[27] S. Prunet and A. Lazarian (1999), astro-ph/9902314.
[28] X. Wang, M. Tegmark, and M. Zaldarriaga (2001), astro-ph/0105091.
[29] D. M. Brink and G. R. Satchler, *Angular Momentum* (Clarendon Press, Oxford, 1993), 3rd ed.
[30] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics (Non-relativistic theory)* (Butterworth-Heinemann, Oxford, 1977), 3rd ed.
[31] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (Word Scientific, Singapore, 1988).
[32] B. D. Wandelt, E. Hivon, and K. M. Górski, Phys. Rev. D **64**, 083003 (2001), astro-ph/0008111.
[33] A. Stuart, J. K. Ord, and S. Arnold, *Kendall's advanced theory of statistics*, vol. 2A (Arnold, London, 1999), 6th ed.
[34] A. O'Hagan, *Kendall's advanced theory of statistics*, vol. 2B (Arnold, London, 1994).
[35] M. Tegmark, Astrophys. J. **519**, 513 (1999), astro-ph/9809001.